

RN [1]	SEQUENCE FROM N.A.	RA Hickey M.J., Brinkman F.S.L., Huffnagle W.O., Kowalik D.J., Lagrou M.,
RP	STRAIN=KIMS / Biovar Mediaevalis;	RA Garber R.L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RC	MEDLINE=22137863; PubMed=12142430;	RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RX	Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liiss P.,	RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA	Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,	RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.,
RA	Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,	RT "Complete genome sequence of <i>Pseudomonas aeruginosa</i> PAO1, an opportunistic pathogen.";
RA	Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,	RT Nature 406:959-964 (2000).
RA	Perry R.D.;	RL EMBL; AE00481; AAG03833.1; -.
RT	"Genome Sequence of <i>Yersinia pestis</i> KIM.";	DR PIR; C83591; C83591.
RL	J. Bacteriol. 184:4601-4611 (2002).	DR GO; GO:0016787; F:hydrolase activity; IEA.
DR	EMBL; AE013696; AAM84520.1; -.	DR GO; GO:0008237; F:metallopeptidase activity; IEA.
DR	GO; GO:0016787; F:hydrolase activity; IEA.	DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR	DR GO; GO:0008237; F:metallopeptidase activity; IEA.	DR InterPro; IPR002933; Peptidase_M20.
DR	DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.	DR Pfam; PF01546; Peptidase_M20; 1.
DR	InterPro; IPR002933; Peptidase_M20.	KW Hydrolase; Complete proteome.
DR	InterPro; IPR001261; Pept_M20_A.B.	SQ SEQUENCE 427 AA; 45265 MW; D610270A3E193691 CRC64;
DR	PFam; PF01546; Peptidase_M20_1.	Query Match 31.2%; Score 639.5; DB 16; Length 427;
DR	PROSITE; PS00758; ARGE_DAPE_CPG2_1; 1.	Best Local Similarity 38.1%; Pred. No. 3.7e-32; Mismatches 54;保守性 54; Mismatches 190; Indels 9; Gaps 3;
KW Hydrolase.	SEQUENCE 431 AA; 45642 MW; 471C95FACECFAAF4 CRC64;	Qy 2 TLQKAQAAARIEKEIRELSRFSAECP-GVTRLTYTPEHAARETLLIAAMKAALSVREDAL 60
SQ	Hydrolase.	Db 11 TQRHIDGQRLWQSLLMDLARLGATAKGGVCRIALSDLDRAQRDLFVQWCEAAAGCTVSVDRV 70
Query Match 31.4%; Score 642.5; DB 16; Length 431; Best Local Similarity 36.6%; Pred. No. 2.4e-32; Mismatches 145;保守性 66; Mismatches 172; Indels 13; Gaps 3;	Qy 61 GNIIGRRREGTDPELPAIAVGSHHFDSVRNGGMFDGTAGVCALEAARVMLENGYVNRPFF 120	
Matches 145; Conservat 66; Mismatches 172; Indels 13; Gaps 3;	Db 71 GNIIFARRPGRNPDLPPVMTGSHIDTQPTGGKFEDGCFGVMAGLEVIRTLNDLGWETEAPLE 130	
Qy 18 LSRFSAEGPGVTRLTYTPEHAARETLLIAAMKAALSVREDALGNIGRREGTDPELPAI 77	Qy 121 FIAIVEEGEGARFSSGMLGGRAIAGLVADRELDSLVDDEGVSVRQAATAFGLKREGELQAAA 180	
Db 25 LAISESPEGLTRYVLSPEHLRANRQVGEWMQAVGMQVWQDTVGNCIGRYEGRQDAPAI 84	Db 131 VVVWTNEEGSRFAPCMGSGVFAKGKFTLEBTLAKRDAAGDGVSGVEALDAIG---YAGA 184	
Qy 78 AVGSHFDSVRNGGMFDGTAGVCALEAARVMLENGYVNRP-----PFEFIAIVEEEGAR 131	Qy 181 RSA--DLRAFIELHIEQGPITLEEQQIEIGVYTSIVGVRALRVAVKGRSDDHACTTPMHLR 238	
Db 85 LLGSHLDTVRNAGRYDGMGLGVLTALLEV-----GYLHRHQQLPVAIEVIGFADEEGTR 138	Db 185 RDCLGHGPVGAYPEAHIEQGPILEDEEKTIGVVLGALGQXWFDSLRLRGVEAHAGGPTPMHLR 244	
Qy 132 FSSGMLGGRAIAGLVADRELDSLVDDEGVSVRQAATAFGLKREGELQAAARSAAIDLRAFIE 191	Qy 192 LHTEQGPITLEEQQIEIGVYTSIVGVRALRVAVKGRSDHAGTTMPHLRQDALVPAALMRE 251	
Db 139 FGITLLGSKGVTGRWPVEWLN-TDADGISVAQAMVTRAGLDPMDIGQSARANAFCAYLE 197	Db 198 LHIEQQGPLLENAGLALGVVTDINGARRLQCQFTGLAGHAGTVPMGQRQDALLAGAAEWCV 257	
Qy 192 LHTEQGPITLEEQQIEIGVYTSIVGVRALRVAVKGRSDHAGTTMPHLRQDALVPAALMRE 251	Qy 252 VNRFVNEIADGTVATVGHLTVAPEGGNQVPGEVDFTLDRSPHEESLRLDRISVMVGE 311	
Db 198 LHIEQQGPLLENAGLALGVVTDINGARRLQCQFTGLAGHAGTVPMGQRQDALLAGAAEWCV 257	Db 258 VEALTAQGEHLVATVGTLLTCLPGAVNVIPGQVRLLDIRGPNDLRGNDLITRLAAEAA 317	
Qy 252 VNRFVNEIADGTVATVGHLTVAPEGGNQVPGEVDFTLDRSPHEESLRLDRISVMVGE 311	Qy 312 VASQAGVAADVDEFNLSPVQLAQPTMVDAVREAASALQFTTHRDISSGAGHDMSMFIAQVTD 371	
Db 258 VEALTAQGEHLVATVGTLLTCLPGAVNVIPGQVRLLDIRGPNDLRGNDLITRLAAEAA 317	Qy 318 IATRRGITEAAEGFYRIKATACDSALQQCISISQVGRCLALPSGAGHDAIAMAECWP 377	
Qy 312 VASQAGVAADVDEFNLSPVQLAQPTMVDAVREAASALQFTTHRDISSGAGHDMSMFIAQVTD 371	Db 365 AGHDAFLAELGPAGMIFVPCENGISHNEIENASPDDLAAGCAVLLRAM 413	
Db 318 IATRRGITEAAEGFYRIKATACDSALQQCISISQVGRCLALPSGAGHDAIAMAECWP 377	Search completed: May 3, 2004, 18:21:58	
Qy 372 VGMVFVPSRAGRSHVPEEWTDDDLRKGTTEVVLRVM 407	Job time : 48 secs	
Db 378 VGMVFVRCGGVSHHPDESVTSSDVAVAIQAYLEAV 413		

RESULT 15

Q91673 PRELIMINARY; ID Q91673 PRT; 427 AA.
 AC Q91673; DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE N-carbamoyl-beta-alanine amidohydrolase.
 GN PA0444.
 OS *Pseudomonas aeruginosa*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; *Pseudomonadales*;
 OC Pseudomonadaceae; *Pseudomonas*.
 OX NCBI_TaxID=287;
 RN [1] SEQUENCE FROM N.A.

RP STRAIN=ATCC 15692 / PAO1;
 RC MEDLINE=20437337; PubMed=10984043;
 RX Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,

OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;	
OC	Phyllobacteriaceae; Mesorhizobium.	
OX	NCBI_TaxID=381;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=MAFF303099; PubMed=11214968;	
RX	MEDLINE=21082930; Pubmed=1158360;	
RA	Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S., Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G., Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L., Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M., Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G., Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V., Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.; "Complete genome structure of the nitrogen-fixing symbiotic bacterium Mesorhizobium loti."; RT "Genome sequence of <i>Yersinia pestis</i> , the causative agent of plague.";	
RL	DNA Res. 7:331-338 (2000).	Nature 413:523-527 (2001).
DR	EMBL; AP002996; BAA8439.1; -.	EMBL; AJ414156; CAC92484.1; -.
DR	GO; GO:0016020; C:membrane; IEA.	DR PIR; AH0394; AH0394.
DR	GO; GO:0016787; F:hydrolase activity; IEA.	GO; GO:0016787; F:hydrolase activity; IEA.
DR	GO; GO:0008237; F:metallopeptidase activity; IEA.	GO; GO:0008237; F:metallopeptidase activity; IEA.
DR	GO; GO:0005215; F:transporter activity; IEA.	GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.	InterPro; IPR002933; Peptidase M20.
DR	GO; GO:0006810; P:transport; IEA.	InterPro; IPR001261; Pept M20_A_B.
DR	InterPro; IPR000515; BPD transp.	Pfam; PF01546; Peptidase M20; 1.
DR	InterPro; IPR002933; Peptidase M20.	DR PROSITE; PS00758; ARGE_DAPE_CPG2_1; 1.
DR	Pfam; PF01546; Peptidase M20; 1.	KW Hydrolase; Hypothetical protein; Complete proteome.
DR	PROSITE; PS00402; BPD_TRANS_P_INN_MEMBER; 1.	SQ SEQUENCE 430 AA; 45510 MW; 9C13B117DED0F466 CRC64;
RW	Hydrolase; Complete proteome.	Query Match 31.4%; Score 642.5; DB 16; Length 430;
SEQUENCE	421 AA; 45172 MW;	Best Local Similarity 36.6%; Pred. No. 2.4e-32;
Best Local Similarity	37.1%; Pred. No. 2.4e-32;	Mismatches 66; Mismatches 172; Indels 13; Gaps 3;
Matches	151; Conservative 68; Mismatches 161; Indels 27; Gaps 6;	
Query	10 RIEKEIRELSRFA-EGPGVTRLTYTPEHAAARETLIAAMKAAPALSVERDALGNITIGRRE 68	QY 18 LSRSFAEGPGVTRLTYTPEHAAARETLIAAMKAAPALSVERDALGNITIGRREGTDPELPAl 77
Db	19 QIQRLLFDLGKFGAHSGTGWRTVSPWEVVAADQFAEWCREAGLAVHSDAVGNVWGRLE 78	Db 24 LAALISESPEGLTRVYLSPPEHLRANRQVGEWMOAVGMQWQDVTGNICGRYGRQPDAPAI 83
QY	69 GTDPELPAIAVGSHFDSVRNGGMFDGTAGVVCALEAARVMLENGYYNRHPFEFIAIVEEE 128	QY 78 AVGSHFDSVRNGGMFDGTAGVVCALEAARVMLENGYYNRHPFEFIAIVEEE 131
Db	79 GEEPG-PSIVSGSHIDSQTGGRYDGAIAALVAVDALAKOFGKPRRTLEVLAICEEE 137	Db 84 LLGSHLDTVRNAGRYDGMVLGVLTAEVV----GYLHRHQQLPVAIEVIGFADEEGTR 137
QY	129 GARFSSGML-GGRAIAGLVADRELDSLVEDDGVSVRQAATAFGKLKPGELOQAARSAAADLR 187	QY 132 FSSGMLGGRAIAGLVADRELDSLVEDDGVSVRQAATAFGKLKPGELOQAARSAAADLRAlFIE 191
Db	138 GSRFPAAASLWGSRAITGCIDPREFDELVDGDGVSIGQAMKAVGLDTNDLGKAQRD--DIG 195	Db 138 FGITLLGSKGVTGRWPVWEWLNT-TDADGISVAQAMVRAGLDPMDIGQSARAANAFCAYLE 196
QY	188 AFIELHIEQGPILQEQQIEQVVTISIVGVRAVKGRSIDHAGTTPMHILQODALVPAALMRE 247	QY 192 LHIEQGPILQEQQIEQVVTISIVGVRAVKGRSIDHAGTTPMHILQODALVPAALMRE 251
Db	196 TFIELHIEQGPVLEAAGLGVGVLTAINGLRHYRVELAGEANHAGAFPMDDRPMAGFA- 254	Db 197 LHIEQGPILEAGLGVVTIDINGARRLQCFQFTGLAGHAGTVPMQRCQDALAGAAEWMCV 256
QY	248 MREVNRFPVNEIADG-----TATVGHLLTVAPGGNNQVPGEVDFTLDRSPHEE 296	QY 252 VNRFVNIEIADGTVATVGHLLTVAPGGNNQVPGEVDFTLDRSPHEE 311
Db	255 -----EIAGGLLISTAERWGRPAVTTVGQSVEPNLPAIIPAKVTFMIDARHPDPD 304	Db 257 VEALTAQGEHLVATVGTLLTCLPGAVNVIPQVRLTLDIRGPNDRGVNDLLTRLLAEAEA 316
QY	297 SLRVLDRISSVWGEVASQAGVAADVDEFNLSPVOLAPTVMDAVREAASALQFTHRDIS 356	QY 312 VASQAGVAADVDEFNLSPVOLAPTVMDAVREAASALQFTHRDISGAGHDSMFIAQVTD 371
Db	305 AVQRLYTLHENLHEVADRRQLKVKITVVENQVPLICHPEIVTAIKATAEQGIRLSNLS 364	Db 317 IATRRGIFTAAEGFYRIKATACDSALQQCISQSISQVQGRCLALPSGAGHDATAIAECWP 376
Qy	357 SGAGHDMSMFIQVQTDVGMVFPSRAGRSHVPEEWTDDDLRKGTTEVLLRM 407	Qy 372 VGMVFVPSRAGRSHVPEEWTDDDLRKGTTEVLLRM 407
Db	365 SGGSHTDQQMSRIARAGRIFVRSKDRSHTPEFFSSIDDIVDGKVL 411	Db 377 VGMFLFVRCKGGVSHHPDESVTSSDVAVAIQAYLEAV 412
		RESULT 14
		Q8D176 PRELIMINARY; PRT; 431 AA.
RESULT 13		ID Q8D176 PRELIMINARY; PRT; 431 AA.
Q8ZBY1		ID Q8D176; AC Q8ZBY1; AC Q8ZBY1; DT 01-MAR-2003 (TREMBLrel. 23, Created)
AC		DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT	01-MAR-2002 (TREMBLrel. 20, Created)	DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)	DE Putative N-carbamyl-L-amino acid amidohydrolase.
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)	GN Y0939.
DE	Putative amino acid hydrolase.	OS Yersinia pestis.
GN	YPO3249.	OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OS	Yersinia pestis.	OC Enterobacteriaceae; Yersinia.
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;	NCBI_TaxID=632;
OC	Enterobacteriaceae; Yersinia.	

OC	Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;	
OC	Pasteurellaceae; Pasteurella.	
OX	NCBI_TaxID=747;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=PM70;	SEQUENCE FROM N.A. / NCPPB 528;
RX	MEDLINE=21145866; PubMed=11248100;	REINACH F.C., Farah C.S., Furian L.R., da Silva A.C.R., Ferro J.A., Reinaldo C.B., Van Sluy M.A., Almeida N.F., Monteiro-Vitorreli C.B., Camargo L.E.A., Alves L.M.C., do Amaral A.M., Bertolini M.C., Chambergo F., Ciapina L.P., Cannarotte G., Cardozo J., Cursino-Santos J.R., El-Dorry H., Ciccarelli R.M.B., Coutinho L.L., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T., Faria J.B., Ferreira A.J.S., Formighieri E.F., Franco M.C., Graggio C.C., Gruber A., Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F., Locali E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M., Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H., Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R., Pereira H.A., Rossi A., Sena J.A.D., Souza R.F., Spinola L.A.F., Takita M.A., Tamura R.E., Tezza E.C., Tezza R.I.D., Trindade dos Santos M., Truffi D., Tsai S.M., White F.F., Setubal J.C., Kitajima J.P.; "Comparison of the genomes of two Xanthomonas pathogens with differing host specificities.", Nature 417:459-463 (2002).
RL	EMBL; AE006050; AAK02244.1;	DR GO; GO:0008237; F: metallopeptidase activity; IEA.
DR	GO; GO:0005622; C: intracellular; IEA.	DR GO; GO:0003735; F: structural constituent of ribosome; IEA.
DR	GO; GO:0006412; P: protein biosynthesis; IEA.	DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.
DR	GO; GO:0006508; P: proteolysis and peptidolysis; IEA.	DR InterPro; IPR002933; Peptidase M20.
DR	InterPro; IPR001261; Pept_M20_A_B.	DR InterPro; IPR002171; Ribosomal_L2.
DR	PFam; PF01546; Peptidase M20; 1.	DR PROSITE; PS000467; RIBOSOMAL_L2.
DR	PROSITE; PS000758; ARGE_DAPE_CPG2_1; 1.	DR PROSITE; PS000467; RIBOSOMAL_L2.
KW	Hypothetical protein; Complete proteome.	
SQ	SEQUENCE 412 AA; 44898 MW; C3CC18258DD05B61 CRC64;	SEQUENCE 412 AA; 44898 MW; 1C58D82452E22FB5 CRC64;
Query Match	32.4%; Score 663; DB 16; Length 412;	Query Match 31.5%; Score 645.5; DB 16; Length 423;
Best Local Similarity	36.1%; Pred. No. 1.e-33;	Best Local Similarity 39.2%; Pred. No. 1.5e-32;
Matches 144; Conservative 74; Mismatches 181; Indels 0; Gaps 0;		Matches 153; Mismatches 171; Indels 9; Gaps 4;
Qy	9 ARIEKEIRELSRFSAEGPGVTRLTYTPEHAAARETLLIAAMKAALSVERDALGNITIGRE 68	Qy 21 FSAEGPGVTRLTYTPEHAAARETLLIAAMKAALSVERDALGNITIGREGTDPELPAPAIAGV 80
Db	6 ARVQSIIDTLASMSSEVEGLTRLAFLSREDEKAHKYIVDLCQTYDLAIRRDEIGNLFIRKA 65	Db 29 FSDPTPTGLFRSWLSPAHRRATTEQVGAWMRQAGMQRVLDAANLVGRYEGAHHAPALLIG 88
Qy	69 GTDPPELPATAVGSHFDSVRNGGMFDGTTAGGVCALEAARVMLENGVNRHPPEFIAVEEE 128	Qy 81 SHFDSVRNGGMFDGTTAGGVCALEAARVMLENGVNRHPF-EFIAIVEEGARFSSGMLG 138
Db	66 GIEDHLPVVAFGSHIDTVNVAGKLDGPLGSVAGLEILFOLCEQQIKTRYPLELIIFTCEE 125	Db 89 SHLDSVRAAGRYDGPLGVLLGIECVAAHLAQG -RRLPFAIEUVFGDEEGSRFPASMFC 146
Qy	129 GARFSSGMLGGRAIALGLVADRELSDLVDEDGVSVRQAATAFGKLKPGELOQAARSAAADLRA 188	Qy 139 GRAIAGLVADRELDSLVDDEDGVSVRQAATAFGKLKPGELOQAARSAAADLRAFIELHIEQGP 198
Db	126 SSRFNYATLGSKVNMCGVVEQAGLSHLRDKQGTAAFAQALADIGLDFSTLHHAKRTAEFFKC 185	Db 147 SRAVAGTL-DAAAALAVRDPDGDVATAAAGWGLDAARLHEAARVPSVLAYLETHIEQGP 205
Qy	189 FIELHIEQPILEQQEIQIGVVTISIVGVRALRVAVKGRSDFHAGTTPMHLRQDALVPAALM 248	Qy 199 ILEQEQUIGVVTISIVGVRALRVAVKGRSDFHAGTTPMHLRQDALVPAALMREVNRFVNE 258
Db	186 FVELHIEQPRLENKTIGVVTGIAAPTRCVVKIQGADHSGATAMHYRHIDALLGGAEI 245	Db 206 VLEVAQLPVGIVTGIAAQRRFRRLRFDGRAGHAGTTMALLRDAAAEALLMIEQIARS 265
Qy	249 VGEVASSQGVAADVDEFNLSPVQLAPTMVDAVREASAALQFTTHRDISSGAGHDMSMFIAQ 368	Qy 259 IADGTVATVGHLTVAPEGGNQVPGEVDFTLDRSPHEESLRVLIDRISVMGEVASOAGV 318
Db	246 ALAVEQQAAIEAGHTATVGNLMAKPGVMMNVPGYCELLVDIRGIHVEARESFTALQQ 305	Db 266 GDDLVATVGKLEVAGAINVPGRVDTCLDVRGDDHRRDAAVAOQIERALEQVVAARGV 325
Qy	309 VTDVGMVFVPSRAGRSHVPEEWTFD DLLRKGTTEVVLRYM 407	Qy 319 ADVDVDEFFNLSPVQLAPTMVDAVREASAALQFTTHRDISSGAGHDMSMFIAQVTDVGMVFVP 378
Db	306 IEQVAQKRGQLQIELQKPSHRSRGINHNPVLLPQEEMVEQIROSAETLGYSYEVMPMSGAGHDAMHMAT 365	Db 326 AIAVEPLQALAASPCAPALIARLTQAVAAQGITPRPLVSQAGHDAVMMAALCBTAMLFVR 385
RESULT 11		Qy 379 SRAGRSHVPEEWTFD DLLRKGTTEVVLRYM 408
Q8PDQ1	ID Q8PDQ1 PRELIMINARY;	Db 386 CAGGISHHPDEHVDPAD---AEVALAVMR 411
AC	Q8PDQ1; Q8PDQ1;	RESULT 12
DT	01-OCT-2002 (TREMBLrel. 22, Created)	Q98LM4
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)	Q98LM4
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)	Q98LM4
DE	N-carbamyl-L-amino acid amidohydrolase.	AC
GN	AMAB OR XCC0284.	DT 01-OCT-2001 (TREMBLrel. 18, Created)
OS	Xanthomonas campestris (pv. campestris).	DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;	DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
OC	Xanthomonadaceae; Xanthomonas.	DE N-carbamyl-L-amino acid amidohydrolase.
OX	NCBI_TaxID=340;	GN MLR0967.
RN	[1]	OS Rhizobium loti (Mesorhizobium loti).

RESULT 8	Q7WJR1	PRELIMINARY;	PRT;	420 AA.	
AC	Q7WJR1;				
DT	01-OCT-2003	(TREMBLrel. 25, Created)			
DT	01-OCT-2003	(TREMBLrel. 25, Last sequence update)			
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)			
DE	N-carbamoyl-L-amino acid amidohydrolase	(EC 3.5.1.-).			
GN	AMAB OR BPP1366.				
OS	Bordetella parapertussis.				
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Alcaligenaceae; Bordetella.				
OX	NCBI_TaxID=518;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=RB50 / ATCC BAA-588;				
RC	MEDLINE=22827954; PubMed=12910271;				
RA	Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N., Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L., Cerdeno-Taraga A.M., Temple L., James K., Harris B., Quail M.A., Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I., Chillingworth T., Collins M., Cronin A., Davis P., Doggett J., Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K., Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C., Rabinowitzsch E., Rutter S., Sanders M., Saunders D., Seeger K., Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K., Unwin L., Whitehead S., Barrell B.G., Maskell D.J.; "Comparative analysis of the genome sequences of Bordetella pertussis, Bordetella parapertussis and Bordetella bronchiseptica.";				
RA	Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I., Chillingworth T., Cerdeno-Taraga A.M., Temple L., James K., Harris B., Quail M.A., Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K., Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C., Rabinowitzsch E., Rutter S., Sanders M., Saunders D., Seeger K., Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K., Unwin L., Whitehead S., Barrell B.G., Maskell D.J.; "Comparative analysis of the genome sequences of Bordetella pertussis, Bordetella parapertussis and Bordetella bronchiseptica.";				
RA	Leather S., Rutter S., Sanders M., Saunders D., Seeger K., Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K., Unwin L., Whitehead S., Barrell B.G., Maskell D.J.; "Comparative analysis of the genome sequences of Bordetella pertussis, Bordetella parapertussis and Bordetella bronchiseptica.";				
RA	DR EMBL; BX640444; CAAE32926.1; -.				
KW	Hydrolase; Complete proteome.				
SQ	SEQUENCE 420 AA; 44863 MW; A8A6911CE3AD870B CRC64;				
Query Match	33.7%; Score 691; DB 16; Length 420;				
Best Local Similarity	37.0%; Pred. No. 2.1e-35;				
Matches	153; Conservative 72; Mismatches 186; Indels 4; Gaps 2;				
QY	2 TLQKAQQARIEKEIRELSRFSAEQPGVTRLTYTPEHAAARETLIAAMKAAALSVR				
DB	61 REDALG 61				
QY	4 TFPPLNAERLMSRVETLSRMTLPDVWTRAFSPFQSRLPFEAGLAVRLDAGG 63				
QY	62 NIIGRREGTDPPELPAIAVGSHFDSVRNGGMFDGTAGVYCALEAARVMLENGVNRHP				
DB	FEF 121				
QY	64 NLVGTRPGRDAARRPIATGSICDTMAGGRFDGIIGVLAGIEVAHTLHEQGVTL				
DB	EHPPFEV 123				
QY	64 NLAGTRPGRDAARRPIATGSICDTMAGGRFDGIIGVLAGIEVAHTLHEQGVTL				
DB	EHPPFEV 123				
QY	122 IAIVEEGARFSSGMLGGRAIAGLVADRELSDLVDEDGVSVRQAATAFGLKPGELQAAAR 181				
DB	124 IDFLSEEPSDYGISCVGSRALSGQLTP-DMLAARNADGETLAEGLRIGGPDALGAPLR 182				
QY	182 SAADLRAFIELHIEEQPILEQEQIEGVVTSIVGVRAVKGRSDDHAGTTPMHLRQDA 241				
DB	183 GPGQTAAFVELHIEEQGPVLESRGGLPIGVVTNIVGTRRVLTVHGQPDHAGTTPMDIRRDA 242				
QY	183 GPGQTAAFVELHIEEQGPVLESRGGLPIGVVTNIVGTRRVLTVHGQPDHAGTTPMDIRRDA 242				
DB	183 GPGQTAAFVELHIEEQGPVLESRGGLPIGVVTNIVGTRRVLTVHGQPDHAGTTPMDIRRDA 242				
QY	242 LVPAALMREVNRNFVNIEIADG--TVATVGHLTVAPEGGNQNVPGEVDFTLDLRSPEESL 298				
DB	242 LVPAALMREVNRNFVNIEIADG--TVATVGHLTVAPEGGNQNVPGEVDFTLDLRSPEESL 298				
QY	243 LVGAARVIDAHHRRASTQSCKPHYYVATIGRIAMTPNAPNAVPGMVLEMLERVSDDAVL 302				
DB	243 LVGAARVIDAHHRRASTQSCKPHYYVATIGRIAMTPNAPNAVPGMVLEMLERVSDDAVL 302				
QY	299 RVLIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMVDAVREAAASALQFTHRDISSG 358				
DB	299 RVLIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMVDAVREAAASALQFTHRDISSG 358				
QY	303 DAFPEAIMAEVADALAALRILSAAAMA PVSRAHPTDCAPLVMDAVEDAAATRLGYASMR				
DB	LPSLG 362				
Q9CP93	RESULT 10				
ID	Q9CP93	PRELIMINARY;	PRT;	412 AA.	
AC	Q9CP93;				
DT	01-JUN-2001 (TREMBLrel. 17, Created)				
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)				
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)				
DE	Hypothetical protein PM0160.				
GN	PM0160.				
OS	Pasteurella multocida.				
RESULT 9	Q7WAL0	PRELIMINARY;	PRT;	420 AA.	
ID	Q7WAL0				
AC	Q7WAL0;				

Db	190	IKAYIEMHIEQKVLEEHDLISIGIVTDIRQPVWLDVTLEGAAADHAGATPMDMRKD---A	245
Qy	246	ALMVRREVNRFVNNEIA--DGTIVATVGHLTIVAPGGGNQVPGEVDFTLDLRSPEESLRVLID	303
Db	246	GLAMAEVLLAVEAISKEHQGVCTVGKMSIEPGGVNIIIPGRACFSVDLRHIRKERQHMVD	305
Qy	304	RISVMVGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREASALQFTTHRDISSGAGHDS	363
Db	306	DLHEQVEAICNQRGVTYNDVKKEVEPATCSHEMVGLIDEVCTELNIRAMKMPGAGHDA	365
	RESULT 6		
	Q8GQGS	PRELIMINARY;	PRT; 409 AA.
ID	Q8GQGS		
AC	Q8GQGS;		
DT	01-MAR-2003	(TREMBLrel. 23; Created)	
DT	01-MAR-2003	(TREMBLrel. 23; Last sequence update)	
DT	01-OCT-2003	(TREMBLrel. 25; Last annotation update)	
DE	N-Carbamoyl-L-amino acid amidohydrolase.		
GN	Geobacillus kaustophilus.		
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.		
OX	NCBI_TaxID=1462;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Chien H.R., Hsu W.;		
RT	"Cloning, expression, and characterization of the thermostable n-carbamoyl-1-amino acid amidohydrolase from Geobacillus kaustophilus ccrcl1223."		
RL	Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL; AP425838; AAN31517.1; -.		
DR	GO:0005634; C:nucleus; IEA.		
DR	GO:0003677; P:DNA binding; IEA.		
DR	GO:0016787; P:hydrolase activity; IEA.		
DR	GO:0008237; F:metallopeptidase activity; IEA.		
DR	GO:0006508; P:proteolysis and peptidolysis; IEA.		
DR	InterPro; IPR001005; Myb_DNA_binding.		
DR	InterPro; IPR002933; Peptidase_M20.		
DR	PFam; PF01546; Peptidase_M20; _1.		
DR	PROSITE; PS00758; ARGE_DAPE_CPG2_1; 1.		
KW	Complete proteome.		
SQ	SEQUENCE 430 AA; 45954 MW; E3BD103B9599E532 CRC64;		
	Query Match 34.7%; Score 710.5; DB 16; Length 430;		
	Best Local Similarity 38.5%; Pred. No. 1.3e-36;		
	Matches 154; Conservative 73; Mismatches 168; Indels 5; Gaps 3;		
Qy	11	IEKEIRELSRFSAEGPGVTRLTYTPEHAAARETLIAAMKAAALSVERDALGNIIGRREGT	70
Db	22	IVRRINVLGAISEETDKLTRIYLSKELRAAADLILGWMREAGMSAHLDAIGNVCGRYEGE	81
Qy	71	DPELPALAVGSHFDSVRNGMFEDGTAGVVCALEAARVMLENGYVNRHPP--EFLIAVEEE	128
Db	82	RPGAPCMLGSHYDTVRDAGKWDGPLGVITAIACVADLNRRG--KRLPFAIEVIGFADEE	139
Qy	129	GARFSSGMIGGRAIAGLVADRELDSLVDEDGVSVRQAATAFGLKPGELQAAARSAAADLRA	188
Db	140	GVRFASTLGSRAVAGTFDESVLNTR-DRDGVSMSRDLVKGFLDPDHVGAALARRELLA	198
Qy	189	FIELHIEQGPITLEQEIQVTVSIVGURALRVAKGRSDHAGTTPMHLRQDALVPAALM	248
Db	199	:YLELHIEQGPVLEAQNLPVGVVTAIAAGATRLAVRLNGMAGHAGTVPMAIRRQDALTGAEC	258
Qy	249	VREVNRFVNEIADGTVAATVGHLTIVAPGGGNQVPGEVDFTLDLRSPEESLRVLIDR	308
Db	259	: :	318
Qy	319	IEAIAKRQLQDVTHEENRTAPCASWLKDQIAQIAAEGVSVFDLPSGAGHDGMAMID	378
Qy	369	VTDVGMVFVPSRAGRSHVPEEWTDFFDLRKTEVV 408	
Db	379	IADVGMIFVRCRGGVSHHPDEHVELADADAGARVLLRVE 418	

Query Match	41.6% ; score 853 ; DB 16; Length 414;	Db	66 EGTRSSGMLGGRAIAGLVADELDLVLVEDGVSVRQAATAFGGLKPGELQAAARSAAADR 1.25
Best Local Similarity	43.2% ; Pred. No. 1.5e-45;	Qy	128 EGARFSSGMLGGRAIAGLVADELDLVLVEDGVSVRQAATAFGGLKPGELQAAARSAAADR 1.87
Matches	173; Conservative 73; Mismatches 152; Indels 2; Gaps 2;	Db	126 EGSRFGGGLMGSRGMGFDGTAGVTRLTYTPEHAAARETTIAAMKAALSVRDEALGNIGRR 68
Qy	10 RIEKEIRELSRFS-A-EGPGVTRLTYTPEHAAARETTIAAMKAALSVRDEALGNIGRR 68	Qy	126 EGSRFGGGLMGSRGMGFDGTAGVTRLTYTPEHAAARETTIAAMKAALSVRDEALGNIGRR 68
Db	7 RIKKHIENLDRFTATPGQGTRLTYSKEDLGARNYLKEEMAKVGLTVSEDAIGNIYGRLE 66	Qy	188 APIELHIEQGPITLEQQEIGVVTISVGVRALRVAVKGRSDHAGTTPMHLRQDALVPAAL 2.47
Qy	69 GTDPELPAIAVGSHFDSVRNGGMFDGTAGVTRLTYTPEHAAARETTIAAMKAALSVRDEALGNIGRR 128	Db	184 SYLELHIEQGPITLEQQEIGVVTISVGVRALRVAVKGRSDHAGTTPMHLRQDALVTAAQ 2.43
Db	67 GDNPDPAPIAVGSHFDSVPNGGAFDGPAGVITGLEVASVFHEQQIKPHFPLEIIAMVEEE 126	Qy	248 MREVNVRFVNIEADGTVATVGHLTVAPGGGNQVPGEVDFTLDLRSPEESLRVLIDRISV 3.07
Qy	129 GARFSSGMLGGRAIAGLVADELDLVLVEDGVSVRQAATAFGGLKPGELQAAARSAAADR 188	Db	244 MIAQIPSLAEEEGEGTVATTGKLNVFPNGSNVIPSKTVFTVDIRSGKEEHIQHVIDKLHE 3.03
Db	127 GSRFGAGLLASRTITGKVTLHEMKDINGVTAEEAMANLGFDDANQVHTAIRSKDSVKA 186	Qy	308 MVGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTA 3.67
Qy	189 FIELHIEQGPITLEQQEIGVVTISVGVRALRVAVKGRSDHAGTTPMHLRQDALVPAAL 248	Db	304 MANSY-NRDGIKITISQQLYMEPKAMNPDIALLKETSSFDIPIYCSMNSGAGHDAMVLA 3.62
Db	187 FIELHIEQGPVLENANEDVALVDTVVGTLTEIKVTVKGQAGHAGTTPMLERKDALSAAVQI 246	Qy	368 QVTDVGVMVFVPSRAGRSHVPEEWTDDLRKGTEVVLRVMKAL 4.10
Qy	249 VREVNRFVNIEADGTVATVGHLTVAPGGGNQVPGEVDFTLDLRSPEESLRVLIDRISV 3.08	Db	363 EVTDVGMLFIPSQDKGSHTPPEEWSDSLIAKAVEILFAAKKL 4.05
Db	247 LNKLPLEAIQEGGGTVLTIKGKLNVYPNGANVIPNKVVFPTVDIRAKDEIHQNTLEKTKV 3.06		
Qy	309 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68		RESULT 5
Db	307 I-QAFERKNGIMCEIEDMLYEKPTHLSSTEIHQALTESADKLGLKYRTMVSGAGHDAMIFAS 365	Q9KET8	PRELIMINARY ; PRT; 4.14 AA.
Qy	369 VTDVGVMVFVPSRAGRSHVPEEWTDDLRKGTEVVLRVMK 4.08	ID	Q9KET8 ; PRELIMINARY ; PRT; 4.14 AA.
Db	366 LTEVGLIFVPSHKGISHAPEEWTDYDKLQKGIEVVLLETVK 4.05	AC	AC 9KET8 ; PRELIMINARY ; PRT; 4.14 AA.
Qy	300 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	AC	AC 9KET8 ; PRELIMINARY ; PRT; 4.14 AA.
Db	308 I-QAFERKNGIMCEIEDMLYEKPTHLSSTEIHQALTESADKLGLKYRTMVSGAGHDAMIFAS 365	DT	DT 01-OCT-2000 (TrEMBLrel. 15, Created)
Qy	367 VTDVGVMVFVPSRAGRSHVPEEWTDDLRKGTEVVLRVMK 4.08	DT	DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
Db	368 LTEVGLIFVPSHKGISHAPEEWTDYDKLQKGIEVVLLETVK 4.05	DT	DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
Qy	301 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	DE	DE N-carbamyl-L-amino acid amidohydrolase.
Db	302 I-QAFERKNGIMCEIEDMLYEKPTHLSSTEIHQALTESADKLGLKYRTMVSGAGHDAMIFAS 365	GN	GN BH0761.
Qy	303 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	OS	OS Bacillus halodurans.
Db	304 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	OC	OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus.
Qy	305 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	OX	NCBI_TaxID=86665; OX NCBI_TaxID=86665;
Db	306 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	RN	RN [1]
Qy	307 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	RP	RP SEQUENCE FROM N.A.
Db	308 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	RC	RC STRAIN=C-125 / JCM 9153;
Qy	309 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	RX	RX MEDLINE=20512582; PubMed=11058132;
Db	310 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	RA	RA Takami H., Nakasone K., Maeno G., Sasaki R., Masui N., Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S., Horikoshi K.;
Qy	311 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	RT	RT "Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis."; RT Nucleic Acids Res. 28:4317-4331 (2000).
Db	312 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	RL	RL EMBL; AP001509; BAB04480.1; - .
Qy	313 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	DR	DR PIR; A83745; A83745.
Db	314 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	DR	DR GO; GO:0016787; F:hydrolase activity; IEA.
Qy	315 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	DR	DR GO; GO:0008237; F:metallopeptidase activity; IEA.
Db	316 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	RA	RA GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
Qy	317 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	RT	RT InterPro; IPR002933; Peptidase_M20.
Db	318 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	DR	DR Pfam; PF01546; Peptidase_M20; 1.
Qy	319 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	KW	KW Hydrolase; Complete proteome.
Db	320 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	SQ	SQ SEQUENCE 414 AA; 45115 MW; 59D9401D81328A11 CRC64;
Qy	321 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	Query Match	Query Match 37.2%; Score 762.5; DB 16; Length 414;
Db	322 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	Best Local Similarity	Best Local Similarity 39.6%; Pred. No. 6.9e-40;
Qy	323 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	Mismatches	Mismatches 161; Conservative 82; Mismatches 153; Indels 11; Gaps 4;
Db	324 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	Matches	Matches 71; KLEGTDLLEPLSWMGSHIDSVPHGGRFDGLVGLGATEAVRTMKEAGIKLKHSLIEVSFT 130
Qy	325 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	QY	QY 10 RIEKEIRELSRFSAEGP---GVTRLTYTPHEAAARETLIAAMKAALSVRDEALGNIG 65
Db	326 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	DB	DB 11 RLWRTLMELAELIGGTTSPIDVHGWTRLSLTKTTELARQVYDLMKDSGLSVQDVAGNIIIG 70
Qy	327 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	QY	QY 66 RREGTDPPELPAIAVGSHFDSVRNGMFDTAGVTRLTYTPHEAAARETLIAAMKAALSVRDEALGNIG 65
Db	328 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	DB	DB 71 KLEGTDLLEPLSWMGSHIDSVPHGGRFDGLVGLGATEAVRTMKEAGIKLKHSLIEVSFT 130
Qy	329 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	QY	QY 126 EEEGARFSGGMLGGRAIAGLVADELDLVLVEDGVSVRQAATAFGGLKPGELQAAARSAAAD 185
Db	330 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	DB	DB 71 DEEGARFAGFIGSGKGAGELTTF-SLADDKGVTYREAFLAANLNPTLYQAIRSDEQ 189
Qy	331 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	QY	QY 186 LRAFIELDPAIAVGSHFDSVRNGMFDTAGVTRLTYTPHEAAARETLIAAMKAALSVRDEALGNIG 65
Qy	332 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	QY	QY 6 ARIEKEIRELSRFS-A-EGPGVTRLTYTPHEAAARETTIAAMKAALSVRDEALGNIGRR 67
Db	333 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	DB	DB 6 ARIEEHDAISKFTATPGQGTRLTYSKDQKARYIKEKMAEYDLESEDFGFGNIFGKL 65
Qy	334 VGEVASQAGVAADVDEFFNLSPVQLAPTMVDAVREAAASALQFTHRDIISSGAGHDMSMFTAQ 3.68	QY	QY 186 LRAFIELDPAIAVGSHFDSVRNGMFDTAGVTRLTYTPHEAAARETLIAAMKAALSVRDEALGNIG 65

Matches	411;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0;	SQ	SEQUENCE	423 AA;	46031 MW;	6333D0DF1C24EBC2 CRC64;
QY	1	MTLQKAQAARIEKEIRELSRFAEGPGVTRLTYPHEAAARETLIAAMKAALSVERDAL	60							Query	Match	41.9%	Score 858;	DB 16;
Db	1	MTLQKAQAERIEKEIRELSRFAEGPGVTRLTYPHEAAARETLIAAMKAALSVERDAL	60								Best Local Similarity	44.2%	Pred. No. 7.3e-46;	
Db	61	GNIIIGREGTDPPELPAIVGSHFDSVRNGGMFDGTAGVVCALEAARVMLENGYVNRPFFE	120							Matches	177;	Conservative	68;	Mismatches 153;
QY	61	GNIIIGREGTDPPELPAIVGSHFDSVRNGGMFDGTAGVVCALEAARVMLENGYVNRPFFE	120							Indels	2;	Gaps	2;	
Db	61	GNIIIGREGTDPPELPAIVGSHFDSVRNGGMFDGTAGVVCALEAARVMLENGYVNRPFFE	120							10 RIEKEIRELSRFS-A-EGPGVTRLTYPHEAAARETLIAAMKAALSVERDALGNIGRRE	68			
QY	121	FIAIVEEGARFSSGMLGGRAIAGLVADRELDLVLVEDDGVSVRQAAATAFGLKPGELQAAA	180							7 RIKKHIEKLDDETFATPGQGTTRLYSKEDDARNYLKGOMAKVGGLTVSEDAIGNIYGRLE	66			
Db	121	FIAIVEEGARFSSGMLGGRAIAGLVADRELDLVLVEDDGVSVRQAAATAFGLKPGELQAAA	180							69 GTDPELPAIVAVGSHFDSVRNGGMFDGTAGVVCALEAARVMLENGYVNRPFFEIAVEEE	128			
QY	181	RSAADLRAPIELHIEQGPITLEQQIEIGVVTSIVGVRALRVAVKGRSDHAGTTPMHLRQD	240							67 GESPDLPAIVVGSHFDSVPNGGAFDGPAGVTGLEVASVFHEQQIKPKHFPLELIAMVEEE	126			
Db	181	RSAADLRAPIELHIEQGPITLEQQIEIGVVTSIVGVRALRVAVKGRSDHAGTTPMHLRQD	240							129 GARFSSGMLGGRAIAGLVADRELDLVLVEDDGVSVRQAAATAFGLKPGELQAAA	188			
QY	241	ALVPAALMVREVNRFVNIEADGTVATVGHLTVAPGGGNQVPGEVDFTLDLRSPHEESLRV	300							127 GSRFGAGLASSRAITGKVTTEMLHEMKDIDGITATDAMAKLGFNANRVRHAIKTESVKA	186			
Db	241	ALVPAALMVREVNRFVNIEADGTVATVGHLTVAPGGGNQVPGEVDFTLDLRSPHEESLRV	300							189 FIELHIEQGPITLEQQIEIGVVTSIVGVRALRVAVKGRSDHAGTTPMHLRQDALVPAALM	248			
QY	301	LIDRISVMVGEVAQSQAGVAADVDEFNLSPVQLAPTMVDAREAAASALQFTHRDIISSGAG	360							187 FIELHIEQGPILENASEDVALVDTVVLQAGHAGTTPMQLDRKDALVSAVEI	246			
Db	301	LIDRISVMVGEVAQSQAGVAADVDEFNLSPVQLAPTMVDAREAAASALQFTHRDIISSGAG	360							249 VREVNRFVNIEIADGTVATVGHLTVAPGGGNQVPGEVDFTLDLRSPHEESLRV	308			
QY	361	HDSMFIAQVTDGMVFVPSRAGRSHYPEWTDDLRKGTEVVLRVMKALDR	412							247 LGQLPPELAIQEGGGTTLVTKLNVYNGANTVIPDKVIFTDIRAKDEIHVQNTLAKTKEI	306			
Db	361	HDSMFIAQVTDGMVFVPSRAGRSHYPEWTDDLRKGTEVVLRVMKALDR	412							309 VGEVASQAGVAADVDEFNLSPVQLAPTMVDAREAAASALQFTHRDIISSGAG	368			
QY	366	LTEVGLIFVPSHNGISHAPEEWTDYDKLQKGIEVVVLKTVK	405							307 I-QATEKNITCEIKDMIYQQPTHSKEIHQALTESADDQLGFKYRTMVSAGHDAMIFAS	365			
Db	366	LTEVGLIFVPSHNGISHAPEEWTDYDKLQKGIEVVVLKTVK	405							369 VTDVGMVFVPSRAGRSHVPEEWTDDLRKGTEVVLRVMK	408			
<hr/>														
RESULT 2														
Q8Y9J1	ID	Q8Y9J1;	PRELIMINARY;	PRT;	423 AA.					RESULT 3				
AC	Q8Y9J1;									Q92EB9	PRELIMINARY;	PRT;	414 AA.	
DT	01-MAR-2002	(TREMBLrel.	20;	Created)						Q92EB9				
DT	01-MAR-2002	(TREMBLrel.	20;	Last sequence update)						AC				
DT	01-JUN-2003	(TREMBLrel.	24;	Last annotation update)						DT	01-DEC-2001	(TREMBLrel.	19;	Created)
DE	Hypothetical protein lmo0537.									DT	01-DEC-2001	(TREMBLrel.	19;	Last sequence update)
GN	lmo0537.									DT	01-JUN-2003	(TREMBLrel.	24;	Last annotation update)
OS	Listeria monocytogenes.									DE	Hypothetical protein lmo0541.			
OC	Bacteria; Firmicutes; Listeriaceae; Listeria.									GN	LIN0541.			
OX	NCBI_TaxID=1639;									OS	Listeria innocua.			
RN	[1]									OC	Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.			
RP	SEQUENCE FROM N.A.									NCBI_TaxID=1642;				
RC	STRAIN=EGD-e / Serovar 1/2a;									RC	SEQUENCE FROM N.A.			
RC	MEDLINE=21537279; PubMed=11679669;									[1]	STRAIN=CLIP 11262 / Serovar 6a;			
RA	Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A., Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T., Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P., Domann E., Dominguez-Bernal G., Duchaude E., Duran L., Dussurget O., Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P., Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D., Jones L.-M., Kaeberl U., Kreft J., Kuhn M., Kunst F., Kurapkat G., Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H., Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R., Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A., Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.; RT "Comparative genomics of Listeria species.";						RA	MEDLINE=21537279; PubMed=11679669;						
RA	RT Science 294: 849-852 (2001).									RA	Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A., Baquero F., Berche P., Bloecker H., Brandt P., Chakraborty T., Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P., Domann E., Dominguez-Bernal G., Duchaude E., Duran L., Dussurget O., Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P., Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D., Jones L.-M., Kaeberl U., Kreft J., Kuhn M., Kunst F., Kurapkat G., Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H., Nordsiek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R., Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A., Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.; RT Science 294: 849-852 (2001).			
RA	RA ListiList; LMO0537;									RA	"Comparative genomics of Listeria species.";			
DR	DR GO; GO:0005622; C:intracellular; IEA.									RA	RT GO; GO:0008237; F:metalloopeptidase activity; IEA.			
DR	DR GO; GO:0008237; F:metalloopeptidase activity; IEA.									RA	DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.			
DR	DR GO; GO:0003700; F:transcription factor activity; IEA.									RA	DR InterPro; IPR002933; Peptidase M20.			
DR	DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.									RA	DR Pfam; PF01546; Peptidase M20; 1.			
DR	DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.									RA	KW Hypothetical protein; Complete proteome.			
DR	DR InterPro; IPR000524; HTH GntR.									RA	KW Sequence 414 AA; 45033 MW; A649E28B34A2A25 CRC64;			
DR	DR InterPro; IPR002933; Peptidase M20.									RA	KW Hypothetical protein; Complete proteome.			
DR	DR PIR; AB1142; AB1142.									RA	KW Hypothetical protein; Complete proteome.			

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 18:17:50 ; Search time 45 Seconds
(without alignments)
2888.744 Million cell updates/sec

Title: US-10-045-063-2

Perfect score: 2049

Sequence: 1 MTLOQAAQARIEKEIRELSR.....FDDLRKGTEVVLRVMKALDR 412

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seq#, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*

1: sp_archea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rabbit:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archeap:*

17 636 31.0 476 10 Q8VXY9
18 631.5 30.8 427 16 Q8BFQ3
19 629 30.7 454 16 Q7UHC4
20 625.5 30.5 425 16 Q8BQ81
21 617.5 30.1 423 16 Q8PQM2
22 612.5 29.9 427 16 Q9A561
23 606 29.6 407 16 Q82ZQ2
24 605.5 29.6 438 16 Q7U3I0
25 604 29.5 423 16 Q7W1K0
26 603.5 29.5 420 16 Q7W7T2
27 601.5 29.4 420 16 Q7WL72
28 600 29.3 423 16 Q7WPJ1
29 599 29.2 414 16 Q98515
30 596.5 29.1 420 16 Q7VWF6
31 591.5 28.9 410 16 Q8ELE3
32 583.5 28.5 416 16 Q98KS4
33 575 28.1 420 2 Q8VUL6
34 563.5 27.5 416 16 Q89R53
35 562 27.4 415 16 Q8UCU8
36 562 27.4 416 16 Q9RV75
37 558.5 27.3 410 16 Q984M3
38 554.5 27.1 414 16 Q92U74
39 548 26.7 416 16 Q92MZ4
40 547.5 26.7 406 16 Q893C9
41 547.5 26.7 415 16 Q7WJ52
42 547.5 26.7 415 16 Q7WA23
43 546.5 26.7 415 16 Q8G2N9
44 545.5 26.6 415 16 Q7VXC4
45 536.5 26.2 415 16 Q8YF79

ALIGNMENTS

RESULT 1			
ID	Q9F464	PRELIMINARY;	PRT; 412 AA.
ID	Q9F464		
AC	Q9F464;		
DT	01-MAR-2001 (TREMBLrel. 16, Created)		
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)		
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)		
DE	L-N-carbamoylase HyuC.		
GN	HYUC.		
OS	Arthrobacter aurescens.		
OC	Bacteria; Actinobacteria; Actinomycetales;		
OC	Micrococcineae; Micrococaceae; Arthrobacter.		
NCBI_TaxID	43663;		
[1]			
RN		SEQUENCE FROM N.A.	
RP		STRAIN=DSM3747;	
RC		MEDLINE=20403588; PubMed=10949312;	
RX		MEDLINE=99210756; PubMed=10194852;	
RA		Wilms B., Wiese A., Syldatk C., Mattes R., Altenbuchner J.,	
RA		Pietzsch M.;	
RA		"Cloning, nucleotide sequence and expression of a new L-N-carbamoylase gene from Arthrobacter aurescens DSM 3747 in E. coli.";	
RT		J. Biotechnol. 68:101-11 (1999).	
RN		[2]	
RN		SEQUENCE FROM N.A.	
RP		STRAIN=DSM3747;	
RC		MEDLINE=20403588; PubMed=10949312;	
RX		MEDLINE=99210756; PubMed=10194852;	
RA		Wiese A., Pietzsch M., Syldatk C., Mattes R., Altenbuchner J.;	
RA		"Hydantoin racemase from Arthrobacter aurescens DSM 3747: heterologous expression, purification and characterization.";	
RT		J. Biotechnol. 80:217-230 (2000).	
RL		EMBL; AF146701; AAC02131.1;	
DR		GO; GO:0008237; F: metallopeptidase activity; IEA.	
DR		InterPro; IPR002933; Peptidase_M20.	
DR		Pfam; PF01546; Peptidase_M20; 1.	
SQ		SEQUENCE 412 AA; 44084 MW; 6EBB76F923FBA2CA CRC64;	
Query Match		99.8%; Score 2044; DB 2;	Length 412;
Best Local Similarity		99.8%; Pred. No. 3e-120;	

Result No.	Score	Query Match	Length	DB ID	Description
1	2044	99.8	412	2 Q9F464	Q9F464 arthrobacte
2	858	41.9	423	16 Q8Y9J1	Q8Y9J1 listeria mo
3	853	41.6	414	16 Q92EB9	Q92eb9 listeria in
4	847	41.3	413	16 QBEKY8	Q8eky8 oceanobacil
5	762.5	37.2	414	16 Q9KET8	Q9ket8 bacillus ha
6	720.5	35.2	409	2 QBGQG5	QBGQG5 geobacillus
7	710.5	34.7	430	16 Q89H48	Q89h48 bradyrhizob
8	691	33.7	420	16 Q7WJR1	Q7wjr1 bordetella
9	687	33.5	420	16 Q7WAL0	Q7walo bordetella
10	663	32.4	412	16 Q9CP93	Q9cp93 pasteurella
11	645.5	31.5	423	16 Q8PDQ1	Q8pdq1 xanthomonas
12	642.5	31.4	421	16 Q98LM4	Q98lm4 rhizobium 1
13	642.5	31.4	430	16 Q8ZBY1	Q8zby1 yersinia pe
14	642.5	31.4	431	16 Q8D176	Q8d176 yersinia pe
15	639.5	31.2	427	16 Q9I673	Q9i673 pseudomonas
16	636	31.0	441	10 Q9FIY0	Q9fiy0 arabidopsis

Search completed: May 3, 2004, 18:21:00
Job time : 21 secs

CC -!- COFACTOR: NADP; contains 2 covalently bound phosphopantetheines.
 CC -!- PATHWAY: Complex polyketide formation in erythromycin biosynthesis.
 CC -!- MISCELLANEOUS: IN EACH ORF OF ERYA TWO MODULES ARE PRESENT EACH ENCODING FOR A FUNCTIONAL SYNTHASE SUBUNIT. THUS ERYA SHOWING 3 ORFS CODES FOR 6 SYNTHASE SUBUNITS. IT IS SUPPOSED THAT EACH SYNTHASE PARTICIPATES IN ONE OF THE SIX FAS-LIKE ELONGATION STEPS REQUIRED FOR FORMATION OF THE POLYKETIDE. MODULE 1, 2, 3, 4, 5, AND 6 PARTICIPATING IN BIOSYNTHESIS STEPS 1, 2, 3, 4, 5, AND 6, RESPECTIVELY.

CC -!- MISCELLANEOUS: BIOSYNTHESIS OF POLYKETIDES; ACYLTRANSFERASE (AT), BETA-KETOACYL CARRIER PROTEIN SYNTHASE (KS), AND ACYL CARRIER PROTEIN (ACP) FOR CHAIN ELONGATION. BETA-KETOREDUCTASE (KR), DEHYDRATASE (DH), AND ENOYL REDUCTASE (ER) FOR PROCESSING OF THE BETA CARBON, AND THIOESTERASE (TE) FOR RELEASE AND LACTONIZATION OF THE FULL-LENGTH CHAIN.

CC -!- SIMILARITY: TO FATTY ACID SYNTHASE (FAS).

CC -!- SIMILARITY: Contains 2 acyl carrier domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR EMBL; X56107; CAA39583.1; -.
 CC DR EMBL; M63677; AAA26495.1; -.
 CC DR EMBL; X62569; CAA44449.1; -.
 CC DR PIR; S13595; SL3595.
 CC DR PIR; S22012; S22012.
 CC DR HSSP; P00101; ICCH.
 CC DR InterPro; IPR001227; Ac_t_trans.
 CC DR InterPro; IPR002198; ADH_short.
 CC DR InterPro; IPR000794; Ketoacyl_synth.
 CC DR InterPro; IPR006163; PP_bind.
 CC DR InterPro; IPR006162; Ppantne_S.
 CC DR InterPro; IPR001031; Thioesterase.
 CC DR Pfam; PF00698; Acyl_transf; 2.
 CC DR Pfam; PF00106; adh_short; 1.
 CC DR Pfam; PF00109; ketoacyl_synt; 2.
 CC DR Pfam; PF02801; ketoacyl_synt_C; 2.
 CC DR Pfam; PF00550; pp-binding; 2.
 CC DR Pfam; PF00975; Thioesterase; 1.
 CC DR PROSITE; PS00012; PHOSPHOPANTETHINE; 2.
 CC DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
 CC DR PROSITE; PS50075; ACP_DOMAIN; 2.
 KW Transferase; Acyltransferase; Antibiotic biosynthesis; NADP; Repeat; Phosphopantetheine; Multifunctional enzyme.
 FT DOMAIN 1 1484 MODULE 5.
 FT DOMAIN 37 3172 MODULE 6.
 FT DOMAIN 37 484 BETA-KETOACYL SYNTHASE 1.
 FT DOMAIN 554 878 ACYLTRANSFERASE (AT) 1.
 FT DOMAIN 1116 1298 BETA-KETOACYL REDUCTASE 1.
 FT DOMAIN 1394 1464 ACYL CARRIER (ACP) 1.
 FT DOMAIN 1488 1954 BETA-KETOACYL SYNTHASE 2.
 FT DOMAIN 2021 2335 ACYLTRANSFERASE (AT) 2.
 FT DOMAIN 2555 2735 BETA-KETOACYL REDUCTASE 2.
 FT DOMAIN 2821 2891 ACYL CARRIER (ACP) 2.
 FT DOMAIN 2926 3172 THIOESTERASE.
 FT ACT_SITE 199 199 ACYL-THIOESTER INTERMEDIATE (BY SIMILARITY).
 FT ACT_SITE 643 643 ACYL-ESTER INTERMEDIATE (BY SIMILARITY).
 FT NP_BIND 1118 1164 NADP.
 FT BINDING 1427 1427 PHOSPHOPANTETHINE (BY SIMILARITY).
 FT ACT_SITE 1661 1661 ACYL-THIOESTER INTERMEDIATE (BY SIMILARITY).
 FT NP_BIND 2557 2605 NADP.
 FT BINDING 2854 2854 PHOSPHOPANTETHINE (BY SIMILARITY).
 FT CONFLICT 231 231 S -> SA (IN REF. 2).
 FT CONFLICT 240 240 MISSING (IN REF. 2).

FT	CONFLICT	289	289	A -> R (IN REF. 2). P -> R (IN REF. 2). PEPRNSLRDTGFTLATRASAMEHRA -> ASRGTRCATPPVS RWPPAAAPWEQ (IN REF. 1).
FT	CONFLICT	493	493	A -> R (IN REF. 2). N -> W (IN REF. 2). E -> D (IN REF. 2). R -> G (IN REF. 2). GNSP -> ARTR (IN REF. 2).
FT	CONFLICT	493	517	R -> G (IN REF. 2). R -> A (IN REF. 2). MISSING (IN REF. 2). AHK -> GIT (IN REF. 2).
FT	CONFLICT	510	510	R -> RQR (IN REF. 2). R -> RELPVYPFQRQR (IN REF. 1). GVAAVPH -> VSLLSRD (IN REF. 2).
FT	CONFLICT	513	513	R -> R (IN REF. 2).
FT	CONFLICT	525	525	MISSING (IN REF. 2).
FT	CONFLICT	536	536	AAH -> RR (IN REF. 2).
FT	CONFLICT	547	551	LCDGRE -> STAER (IN REF. 2).
FT	CONFLICT	553	553	MISSING (IN REF. 1).
FT	CONFLICT	673	673	L -> V (IN REF. 2). A -> R (IN REF. 2).
FT	CONFLICT	716	716	MISSING (IN REF. 2).
FT	CONFLICT	734	736	AA -> RR (IN REF. 2).
FT	CONFLICT	796	896	LCGDRE -> STAER (IN REF. 2).
FT	CONFLICT	896	896	MISSING (IN REF. 1).
FT	CONFLICT	988	994	GVAAVPH -> VSLLSRD (IN REF. 2).
FT	CONFLICT	1108	1116	RTHPLEPLA -> ARTWSPR (IN REF. 2).
FT	CONFLICT	1124	1126	MISSING (IN REF. 1).
FT	CONFLICT	1132	1132	L -> V (IN REF. 2).
FT	CONFLICT	1192	1192	A -> R (IN REF. 2).
FT	CONFLICT	1194	1194	MISSING (IN REF. 2).
FT	CONFLICT	1277	1278	AA -> RR (IN REF. 2).
FT	CONFLICT	1385	1390	LCGDRE -> STAER (IN REF. 2).
FT	CONFLICT	1485	1485	MISSING (IN REF. 2).
FT	CONFLICT	1518	1518	G -> R (IN REF. 2).
FT	CONFLICT	1601	1601	V -> L (IN REF. 2).
FT	CONFLICT	1724	1725	LP -> FA (IN REF. 2).
FT	CONFLICT	1732	1732	Q -> L (IN REF. 2).
FT	CONFLICT	1739	1743	GPAG -> ARRA (IN REF. 2).
FT	CONFLICT	1762	1762	T -> S (IN REF. 2).
FT	CONFLICT	2252	2252	D -> DGAD (IN REF. 2).
FT	CONFLICT	2275	2277	QSP -> AVA (IN REF. 2).
FT	CONFLICT	2408	2408	G -> GR (IN REF. 2).
FT	CONFLICT	2420	2421	LA -> S (IN REF. 2).
FT	CONFLICT	2443	2444	NA -> TH (IN REF. 2).
FT	CONFLICT	2596	2596	A -> G (IN REF. 2).
FT	CONFLICT	2609	2609	P -> A (IN REF. 2).
FT	CONFLICT	2715	2722	RRAEGRAA -> AVRKA VRR (IN REF. 1).
FT	CONFLICT	2754	2754	D -> E (IN REF. 2).
SQ	SEQUENCE	3172	AA;	DBBD5094E77DD5F CRC64; MW;
Query Match				5.5%;
Best Local Similarity				Score 113.5;
Matches 90;	Conservative	21.7%;	Pred. No. 23;	DB 1;
		38;	Mismatches 145;	Gaps 17;
QY				11 IEKEIRELSRFSAECPGVTRLTYTPHEAAARETLIAAMKAALSVREDALGNI ---IGR 66
Db				942 VPSELSDAIRSGLEQSQTAVLTCDVE ---SRSTIGTALEAA ---DTDALSTVGVAAPVH 994
QY				67 REGTDPELPAIAVGSHFDSVRNNGMFDGTAGVVCALLEARVMLENIGVNVRHPFFEFIAIE 126
Db				995 GEAVDPSLDALALVQLA ---GAAGVEAPL ---WULTRNA ---VQVAD 1032
QY				127 EEGARFSSGMGLGGRIAAGLVADRE ---LDSLVDEDGVSVRQAATAFGLKPGELQAA 179
Db				1033 GELVDPQAQMVG ---LGRVVGIEOPGRWGGLVLDVADAASIRSLLAVLADPRGEEQVA 1089
QY				180 ARSAADLRAFIELHIEQPILEQQEIGVVTISVGRLRVAVKGRSDHAGTTPMHLRQ 239
Db				1090 IR -AD ---GIKVARLV ---PAPARART 1109
QY				240 DALVPAALMVREVNRVFVNIEADGTVATVG ---HLTVAPGGGNQV 280
Db				1110 HPLEPLA ---GTVLVTGGGGIGAHLARWLARSGAEHVLGRRGADA 1154
QY				281 PGEVDFTLDRSPHEESLRVLIDRISVMGEVASQAGVAADVDEFNLSPVQLAPTMVDA 340
Db				1155 PG ---ASELREELTALGTGVTIAACDVADRLEAVL ---AAEAAAEGRTVSA 1201
QY				341 VREAASALQFTHRDIISSGAGHDMSMFIAQVTDVGVMFVPSRAGRSHVPEEWTFDFD 394
Db				1202 VMHAAG ---VSTSTPLDDLTEAEFTIADVKV ---RGTVNLDDELCPDLD 1244

Db	69 LGGSTGLGHTRWATHGRPTDNRNAHPHRDAAGKIAVHNGIIFENFAPLRAELEAAAGV---	124	Qy	3 IOKAQQAARIKEK-----IRELSRSFAE-GPGVTRLTYTPEHAAARETLIA 46
Qy	112 GYVNRRHPEFFIAIVEEEGA----RFSSGMLGGRAIAIGLVADE-----LDSLVD 157		Db	315 LQSPKTTLRERERIAFATGIRAPAKYTTRRELIRLEAEMGSRAIWLSRRSSHGVRKEVLEA 374
Db	125 -----EPASDTEAVHLVARQYTGDTAGDFPASVLQLRLEGHTFLVFASADD 176		Qy	47 AM-KAAALSV-VRDAGLNIIGRREGTDPELPAIVGSHFDSVRNNGMFDGTAGUVCALEA 104
Qy	158 DGVSV---RQAATAFGKLKPGELOQAARSAAIDLRAFIELHIEQQPILEQEIGVVTSIV 214		Db	375 AFSPRHSLSDEQKTAIEHVAG----AERIAAVIG----RAG----AGKTTMMKA 416
Db	177 PGTIVARRSTPLVLGIGDGMFV---GSDVAAFIE-HTRDAVELGQDO---AVVLTAD 228		Qy	105 ARVMLENGYVNRHPPFEFFIAVEEGARFSSGMLGGRALAGL----VADRELLS---- 153
Qy	215 GVRA-----LRAVKGRSDHAGTTPMHRLRDALVPAALM--- 248		Db	417 AR-----EAWEAAGYRVVGGALAGKAAEGLKEAGIASRTLSSWELRWN 460
Db	229 GYRITDAGNDHLEAGRDFREFHIDWDLNAAEKGGYDY----FMLKEIAEQPSAVALTL 283		Qy	154 -----LVDEDG-VSVRQ-----AATAFG----LKPGELQAAARSA 183
Qy	249 -----VREVNRFVNEIADGTVATVGHLT-VAPCGGNQNPGEVDFT 287		Db	461 EGRKQOLDDDKTIFVLDEAGMVSSRQMALEFVETATKAGAKLVLVGDPHQPIEAGAAFRAI 520
Db	284 LGHFDKNRIVLDQELREIDK-VFIVACGTAYHSGLLARYAIEHWTRLPEVELA 342		Qy	184 ADLRAFIELHI----EQGPILEQQEIQIGVVTTSIVGV----RALRVAVKGRS---- 227
Qy	288 LDL--RSPHEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMV----D 339		Db	521 ADRIGYAELETIYRQEQWMCDASLDLARGNVGKUVDTYRANGRMMRSELKAEVQNLIA 580
Db	343 SEFRYRDP----VLDRSTLVI--AISQSGETADTLEAVRAHKTKQAKVLAICNTNGS 393		Qy	228 -----DHAGTTPM--HLRQDA-LVPAALMVRVNRFVNEIADGTVATVGHLTVAPGGGN 278
Qy	340 AVREAAASALQFTTHDISSGAGHDMSMFIAQVTD---VGMVFVPSRAGRSHVPEEWTFDDDL 396		Db	581 DWDRDYDPTKTLILAHLLRRDVMLNQMARAKLVERGIIVDAGESFKADEGDNRNFAP--GD 638
Db	394 QIPREADAVLYTRAGPEIGVAAATKTFLQIAANYLVGLALAQAR-GTKYPDEVAREYREL 452		Qy	279 QV----PGEVDFTLDDLSRSPHEESLRLVLIIDRISVMVGEVASOAGVAADVDEFNLSPVQL 333
Qy	397 RKGTENVLRVMKALD 411		Db	639 QIVFLKNEGALGVKNGMRGKVVEAAQ---NRIVAEIGEVEHRRQVMVE-SRFYNNLHDGY 694
Db	453 EAMPDLIKRVLAGMD 467		Qy	334 APTM-----VDAVREAAS 346
			Db	695 ATTIIHKSQGATVDRVKYLAS 714
RESULT 14				
TRAA_RHISN	RHISN	STANDARD;	PRT;	1102 AA.
ID	TRAA_RHISN	STANDARD;	PRT;	3172 AA.
AC	P55418;	ERY3_SACER	STANDARD;	
DT	01-NOV-1997 (Rel. 35, Created)	ID	ERY3_SACER	
DT	01-NOV-1997 (Rel. 35, Last sequence update)	AC	Q54097; Q99270;	
DT	10-OCT-2003 (Rel. 42, Last annotation update)	DT	01-OCT-1993 (Rel. 27, Created)	
DE	Probable conjugal transfer protein traA.	DT	01-NOV-1997 (Rel. 35, Last sequence update)	
GN	TRA A OR Y4DS.	DT	10-OCT-2003 (Rel. 42, Last annotation update)	
OS	Rhizobium sp. (strain NGR234).	DE	01-OCT-2003 (Rel. 42, Last annotation update)	
OG	Plasmid sym pNGR234a.	DE	01-OCT-2003 (Rel. 42, Last annotation update)	
OC	Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;	GN	DEoxyerythronolide B synthase III (DEBS 3).	
OC	Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.	OS	Saccharopolyspora erythraea (Streptomyces erythraeus).	
OX	NCBI_TaxID=394;	OC	Bacteria; Actinobacteria; Pseudonocardiae; Actinomycetales;	
RN	[1]	OC	Pseudonocardineae; Actinomycetaceae; Saccharopolyspora.	
RP	SEQUENCE FROM N.A.	OX	NCBI_TaxID=1836;	
RX	MEDLINE=97305956; PubMed=9163424;	RN	[1]	SEQUENCE FROM N.A.
RA	Freiberg C.A., Fellay R., Bairach A., Broughton W.J., Rosenthal A.,	RP		SEQUENCE FROM N.A.
RA	Perret X.;	RC		SEQUENCE FROM N.A.
RT	"Molecular basis of symbiosis between Rhizobium and legumes.";	RX		SEQUENCE FROM N.A.
RL	Nature 387:394-401(1997).	RA		SEQUENCE FROM N.A.
CC	-!- SIMILARITY: Belongs to the mobA/mobL family.	RT		SEQUENCE FROM N.A.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	RT		SEQUENCE FROM N.A.
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	RN		SEQUENCE FROM N.A.
CC	the European Bioinformatics Institute. There are no restrictions on its	RX		SEQUENCE FROM N.A.
CC	use by non-profit institutions as long as its content is in no way	RA		SEQUENCE FROM N.A.
CC	modified and this statement is not removed. Usage by and for commercial	RT		SEQUENCE FROM N.A.
CC	entities requires a license agreement (See http://www.ibb-sib.ch/announce/	RT		SEQUENCE FROM N.A.
CC	or send an email to license@ibb-sib.ch).	RL		SEQUENCE FROM N.A.
DR	EMBL; ABE000069; AAB91648.1;	RN		SEQUENCE FROM N.A.
DR	InterPro; IPR005053; MobA_MobL.	RP		SEQUENCE FROM N.A.
DR	Pfam; PF03389; MobA_MobL_1.	RC		SEQUENCE FROM N.A.
KW	Conjugation; ATP-binding; Plasmid.	RX		SEQUENCE FROM N.A.
FT	NP BIND 404 411 ATP (POTENTIAL).	RA		SEQUENCE FROM N.A.
SQ	SEQUENCE 1102 AA; 123033 MW; EC17A5528F7B9CAF CRC64;	RT		SEQUENCE FROM N.A.
Query Match	5.6%; Score 114; DB 1; Length 1102;	RT		SEQUENCE FROM N.A.
Best Local Similarity	24.1%; Pred. No. 6.5;	RT		SEQUENCE FROM N.A.
Matches 106; Conservative 48; Mismatches 150; Indels 136; Gaps 23;	RL			SEQUENCE FROM N.A.
CC	CC	CC		SEQUENCE FROM N.A.

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to licensee@isb-sib.ch).

408 LSKDLEAIKKEIR 420

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch) .	
Db 408 LSKDLEAIKKEIR 420	RESULT 13
EMBL; AE001514; AAD06431.1; - .	GLMS_MYCSM
PIR; H71880; H71880.	ID GLMS MYCSM STANDARD ; PRT; 627 AA.
HSSP; P00968; 1CS0.	AC 068956;
HAMAP; MF_01210; - ; 1.	AC 068956; (Rel. 39, Created)
InterPro; IPR006275; Cara_L_glu.	DT 30-MAY-2000 (Rel. 39, Last sequence update)
InterPro; IPR005483; CPase_L.	DT 30-MAY-2000 (Rel. 39, Last annotation update)
InterPro; IPR005479; CPase_L_D2.	DT 28-FEB-2003 (Rel. 41, Last annotation update)
InterPro; IPR005480; CPase_L_D3.	DE Glucosamine--fructose-6-phosphate aminotransferase [isomerizing]
InterPro; IPR005481; CPase_L_N.	DE (EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-phosphate amidotransferase) (GFAT) (L-glutamine-D-fructose-6-phosphate amidotransferase) (Glucosamine-6-phosphate synthase).
InterPro; IPR004362; MGS_like.	DE
Pfam; PF00289; CPSase_L_chain; 2.	GN
Pfam; PF02786; CPSase_L_D2; 3.	OS Mycobacterium smegmatis.
Pfam; PF02787; CPSase_L_D3; 1.	OC Bacteria; Actinobacteridae; Actinomycetales;
Pfam; PF02142; MGS; 1.	OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
PRINTS; PRO0098; CPSASE.	NCBI_TaxID=1772;
TIGRFAMS; TIGR01369; CPSaseII_lrg; 1.	RN
PROSITE; PS00866; CPSASE_1; FALSE_NEG.	RP SEQUENCE FROM N.A.
PROSITE; PS00867; CPSASE_2; 2.	RC STRAIN=ATCC 700084 / mc(2)155;
Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat; ATP-binding; Manganese; Complete proteome.	RA Schaeffer M.L.; Besra G.S.; Belisle J.T.; Inamine J.M.;
ATP-binding; Manganese; Complete proteome.	RT "Biochemical and genetic definition of effects of mannosamine on mycobacteria.";
DOMAIN 1 399	RT Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DOMAIN 400 552	RL CC -!- FUNCTION: Catalyzes the first step in hexosamine metabolism, converting fructose-6P into glucosamine-6P using glutamine as a nitrogen source (By similarity).
DOMAIN 553 951	CC -!- CATALYTIC ACTIVITY: L-glutamine + D-fructose 6-phosphate = L-glutamate + D-glucosamine 6-phosphate.
DOMAIN 952 1085	CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
REPEAT 1 552	CC -!- SIMILARITY: In the C-terminal section; belongs to the SIS family.
REPEAT 553 1085	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation.
NP_BIND 151 208	CC the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
NP_BIND 301 350	CC
METAL 283 283	CC
METAL 297 297	CC
METAL 299 299	CC
METAL 830 830	CC
METAL 842 842	CC
METAL 1085 AA; 119965 MW; 36DEFD313BE5C387 CRC64;	CC
SEQUENCE 1085 AA; 119965 MW; 36DEFD313BE5C387 CRC64;	CC
Y 26 PGVTRLTY----TPEHAAARETLIAAKAAALSVERDALGNIIGRREGTDPELPALAVGS 81	CC
b 58 PEFSHQTYIQPITPENIA-----	DR AAC14295.1; - .
b 82 HFDSV---RNNGMFDGTAGVVCALEAARVMLENGVNRHPFEFIAIVEEGARFSSGML-	DR HSSP; P17169; 1MOS.
b 95 ALNAVQMHQGMLEGVELLGAKIEA----IKKG-----	DR MEROPS; C44.001; - .
Y 138 -GGRAJAGLVADRELDSL--VDEDG--VSVRQAATAFGLKPG-----ELQAAARSAAIDL 186	DR HAMAP; MF_00164; - ; 1.
b 137 IGMDLPKGRYAYSELEAISEIGPAIIIRASFTLAGGGSGVAYNIEEFQELAKNALDA 196	DR InterPro; IPR000583; GATase_2.
Y 187 RAFIELHIEQGPI---LEQEQIEIG----VVTISIVGVRALRVAKGRSDHAGTTPMHL 237	DR InterPro; IPR001347; SIS.
b 197 SPINEILIEESLLGWKEYEMEVIRDGKDNCIIIVCCIENDPMGV-----HTG-----	DR Pfam; PF00310; GATase_2; 1.
Y 238 RODALVPA---ALMVERVNRFVNEIADGTVAATGVHLTVAPPGGNNQVPGEVDFTDLRSRH 294	DR TIGRFAMS; TIGR01135; g1mS; 1.
b 244 -DSITIAPSLLTIDKEYQR---MRDASFAILREIGVDTGGSN----VQFAI----H 287	DR PROSITE; PS00443; GATASE_TYPE_II; 1.
Y 295 EESLRVVL-----IDRISVM-----VGEVASOQGVAAADVDEFFN---LSPVQLAPT 336	KW Transferase; Aminotransferase; Glutamine amidotransferase.
b 288 PETLRMVVITEMPRVRSSALASKATGFIPIAKVATLAVGFLSLDEIKNDITNTPASPEPS 347	FT INIT_MET 0 0 BY SIMILARITY.
Y 337 MVDAVREAASALQFTHRDIISSGAGHSIMFIAQVTDVGMVFPVS-RAGRSHVPEWTDFDD 395	FT ACT_SITE 1 1 GATASE (BY SIMILARITY).
b 348 LDYIVVKIPRFAFEKFGAGVSSTLGTSMKSIGEVMAIGGNFLEALQKALCSLENNWLGFES 407	FT DOMAIN 1 192 ISOMERIZATION FRU-6P (BY SIMILARITY).
Y 396 LRKGTEVVLRVMK 408	SQ SEQUENCE 627 AA; 67771 MW; AA82F590871F926E CRC64; GLUTAMINE AMIDOTRANSFERASE.
Query Match 5.6%; Score 114; DB 1; Length 627;	Query Match 5.6%; Score 114; DB 1; Length 627;
Best Local Similarity 20.2%; Pred. No. 3.5;	Best Local Similarity 20.2%; Pred. No. 3.5;
Matches 100; Conservative 72; Mismatches 167; Indels 156; Gaps 2	Matches 100; Conservative 72; Mismatches 167; Indels 156; Gaps 2
QY 37 HAAARETLIAAK-----AAALSVREDALGNIIGRREGTDPELPALAV-----	QY 37 HAAARETLIAAK-----AAALSVREDALGNIIGRREGTDPELPALAV-----
Db 9 HRPARDIVVDAIRRMEYRGYDSAGGIALDNGNGLTVRRAGLANLEATLAEETDSNDGDG 68	Db 9 HRPARDIVVDAIRRMEYRGYDSAGGIALDNGNGLTVRRAGLANLEATLAEETDSNDGDG 68
QY 80 --GS-----HFD-----SVRNGGMFDGTAGVVCALEAARVMLEN 11	QY 80 --GS-----HFD-----SVRNGGMFDGTAGVVCALEAARVMLEN 11

RESULT 10					
GLMS_MYCLE		STANDARD;	PRT;	624 AA.	
ID	GLMS MYCLE				
AC	P40831;				
DT	01-FEB-1995 (Rel. 31, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	28-FEB-2003 (Rel. 41, Last annotation update)				
DE	Glucosamine--fructose-6-phosphate aminotransferase [isomerizing]				
DE	(EC 2.6.1.16) (Hexosephosphate aminotransferase) (D-fructose-6-				
DE	phosphate amidotransferase) (GFAT) (L-glutamine-D-fructose-6-phosphate				
DE	amidotransferase) (Glucosamine-6-phosphate synthase).				
GN	GLMS OR MU0371 OR B229_C3_238.				
OS	Mycobacterium leprae.				
OC	Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales;				
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.				
OX	NCBI_TaxID=1769;				
RN	[1]				

QY	280	-VPGEVDFITDLRSPHEESLRLVIDRISVMVGEVASQAGVAADVDEFNLSP---VQLAP	335
Db	237	IVCDEVHILAEARSLEVPEKMEAQVKMKAFAFEAAADMGGRAEV-EIEVMYPGFKYQDGD	295
QY	336	TMVDAVREAAASALQFTHRDISSGAGHDMSMFIAQVTDVGMOVFPVPSRAG--RSHVPEEWTFDF	393
Db	296	QVVEIAKKAAAKIGRPSELQTSGGSDANVIA---GHGIPPTVNLAvgYEQIHTKNEKMPI	352
QY	394	DDLRKGTEVVLRVMK	408
Db	353	EELVKTAAEMVVAIE	367
QY	129	GARFSSGMLGGRAI-AGLVADRELDSLVDEDGVSVRQAATAFGGLKPGELQAAARSAAADLR	187
Db	98	GR--SRAHLGGRSVPVGMIAEL-ADELVAVHGQTDQQGLLKLN--RQRQALDRYAGDAV	151
QY	188	AFILEHIEQGPITLEQEIQIEIGVVTTSIVGVRALRVAVKGRSDHAGTTPMLRQDALV-PAA	246
Db	152	A-----GPLAKYAAYRRRLRAVRELEETTRERAQEADLLRYGLDEIAAVEPRA	203
QY	247	----LMVREVNRNFVNIEADGTATVGHILTVAPEGGNQVPGEVD-----FTLDLRS	293
Db	204	GEDVELAEEAERLGHAEAALASAATVAAHALA-GNPEPDPEGVDGATLVAGAQRALDAVRS	261
QY	294	HEESLRVLIDRI---SVMVGEVASQ-AGVAADVDEFNLSPVQLAPTMVDVAREASALQ	349
Db	262	HDPALALAERGILRDVAGELAGYADDLD---ADPLRLA-----AVEERRAALT	312
QY	350	FTHR---DISSGAGHDMSMFIAQVTDVGMOVFPVPSRAGSHVPEEWTDFDDLRKG	399
Db	313	347 ALTRKYGEDIAAVLWSAEQSAARLTLE-----DGDDERIG	347
OS	RESULT 8		
RECN_STRCO	STANDARD;		
ID_Q9S220;	PRT; 572 AA.		
AC	ID_YC78_MYCTU STANDARD;		
RP	ID_Q11042; PRT; 875 AA.		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=A3 (2) / M145;		
RX	MEDLINE=21996410; PubMed=12000953;		
RA	Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,		
RA	Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,		
RA	Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,		
RA	Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,		
RA	Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,		
RA	Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,		
RA	Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,		
RA	Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,		
RA	Hopwood D.A.; "Complete genome sequence of the model actinomycete Streptomyces coelicolor A3 (2)." Nature 417:141-147 (2002).		
RT	RT -!- FUNCTION: MAY BE INVOLVED IN RECOMBINATIONAL REPAIR OF DAMAGED DNA (BY SIMILARITY).		
CC	CC -!- SIMILARITY: Belongs to the recN family.		
CC	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	CC -----		
CC	CC EMBL; AL939110; CAB52844.1; -.		
DR	DR PIR; T36883; T36883.		
DR	DR InterPro; IPR003439; ABC transporter.		
DR	DR InterPro; IPR004604; RecN.		
DR	DR TIGRFAMs; TIGR00634; recN; 1.		
KW	KW DNA repair; ATP-binding; Complete proteome.		
NP BIND	NP BIND 25 32 ATP (POTENTIAL).		
FT	FT SEQUENCE 572 AA; 59836 MW; 1D303A47D50BBC14 CRC64;		
SQ	SQ Query Match 6.2%; Score 128; DB 1; Length 572;		
CC	CC Best Local Similarity 23.7%; Pred. No. 0.39; Matches 98; Conservative 58; Mismatches 158; Indels 100; Gaps 21;		
QY	11	IEKEIRELRSRFSAGPGVTRLYTPEHAARETLIAAMKAALSVERDALGN1IGRREGT	70
Db	9	IDDAVELS-----PGFTAVT---GETGAGKTWVT-----SLGLLGGR--	45
QY	71	DPELPAIAGVSHFDSVRNGGMFDGTAGWVCALEARVMLENG--YVNRPFFEFIAVEEE	128

DR EMBL: U89279; AAB93857.1; -. DR PIR; C64783; C64783.

DR EcoGene; EG13623; a11C.

DR InterPro; IPR002933; Peptidase M20.

DR Pfam; PF01546; Peptidase M20; 1.

KW Hydrolase; Purine metabolism; Complete proteome.

SQ SEQUENCE 411 AA; 45694 MW; C55D1EF854A8F513 CRC64;

Query Match 25.6%; Score 525.5; DB 1; Length 411; Best Local Similarity 32.7%; Pred. No. 7.7e-27; Matches 131; Conservative 63; Mismatches 204; Indels 3; Gaps 2;

Qy 11 IEKEIRELSRSFSAE-GPGVTRLTYTPBHAARETLIAAMKAALSVREDALGNTIGREG 69

Db 9 IEETLIPWLSSFGADPAGGMTRLLYSPWEWLETOQQFFKRMMAASGLETRFDEVGNLYGRNLNG 68

Qy 70 TDPELPAPAIAVGSHFDSVRNGMFDGTAGVVCALEAARVMLENGYVNRPFPFIAIVEEG 129

Db 69 TEYPOQEVLVGSKHNIDTVNGNLDQFGALAAWLAIDWLKTOYGAPLRTVEVVAAMAEEEG 128

Qy 130 ARFSSGMLGGRAIAGLVADRELDLVLVEDGVSVRQAATAFGLKPGELOAARSAAADLRAF 189

Db 129 SRFPYVFWGSKNIFGLANPDPVRNICDAKGNSFVDAMKACGFTLPNAPLTPRQ-DIKAF 186

Qy 190 TELHIEQGPITLEEQEQTEIGVYTSIVGVRAVLRAVKGRSDHAGTTPMHLRODALVPAALMV 249

Db 187 VELHIEQGCVLENSNGQSIGVNAIVGQRRTVTLNGESNHAGTTPMGYRRDTVYAFSRIC 246

Qy 250 REVNRVFVNEIADGTVATVGHTLTAVPGCCNQVPGEVDFTLDLRSPHEESLRVLIDRISVMV 309

Db 247 HQSVEKAKRMGDPVLTFGKVEPRPNTVNVPGKTTFTIDCRHTDAAVLRFDTQQLENDM 306

Qy 310 GEVASOAGVAAADVDEFFNLSPVQLAPTMVDAVREASALQFTHDIISSGAGHDSMFIAQV 369

Db 307 RAICDEMDIGIDIDLWDEEPVPMNKELNATLTELCEERKLNYRVMSMSGAGHDAQIFFAPR 366

Qy 370 TDVGMVFVPSRAGRSHVPEEWTDDLRKGTEVVRLVMKAL 410

Db 367 VPTCMIFIPSINGISHNPAERTNITDLAGVKTLLMLYQL 407

RESULT 7 YQJE_BACSU STANDARD; PRT; 371 AA.

ID YQJE_BACSU

AC P54542;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Hypothetical protein Yqje.

GN YQJE OR BSU23910.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1] SEQUENCE FROM N.A.

RC STRAIN=168 / JH642;

RX MEDLINE=97124195; PubMed=8969508;

RA Mizuno M., Masuda S., Takeuchi M., Hobono S., Sato T., Takeuchi M., Kobayashi Y.;

RT "Systematic sequencing of the 283 kb 210 degrees-232 degrees region of the Bacillus subtilis genome containing the skin element and many sporulation genes.";

RL Microbiology 142:3103-3111 (1996).

RN [2] SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Brignell S.C., Bron S., Borriss R., Bourrier L., Brans A., Braun M., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoff A., Ehrlich S.D., Emmerson P.T.,

Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Galleron N., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Nonne D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Presecan E., Pujić P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schröter R., Scoffone F., Soldo B., Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Sorokin A., Tacconi E., Takagi T., Takaishi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandebol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.; RT "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis"; RT Nature 390:249-256 (1997).

RL FUNCTION: COULD BE A PEPTIDASE.

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).

CC -!- SIMILARITY: Belongs to peptidase family M20A.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMB

RA	Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Presecan E., Pujić P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y., Sato T., Scanlan E., Schleicher S., Schröter R., Scorrone F., Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandembol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Weitzenerger T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshioka H.F., Zumstein E., Yoshikawa H., Danchin A.; "The complete genome sequence of the Gram-positive bacterium <i>Bacillus subtilis</i> ."; <i>Nature</i> 390:249-256 (1997).	Qy	202 QEQIEIGVVTIVGVRLRVAVKGRSDHAGTTPMHLRQDALVPAALMVREVNRFVNNEIAD 2611	RA
RA	[2] FUNCTION. STRAIN=168; MEDLINE=21242727; PubMed=11344136;	RP	202 MSGRDIGVTTIAGGRRYLTLEGECNHAAGTTSMKWKRKDPLAASSRIIHLLRSDELPD 2612	RA
RA	Schultz A.C., Nygaard P., Saxild H.H.; "Functional analysis of 14 genes that constitute the purine catabolic pathway in <i>Bacillus subtilis</i> and evidence for a novel regulator controlled by the PUCR transcription activator."; <i>J. Bacteriol.</i> 183:3293-3302 (2001).	RC	202 GTATVGHLTVAPGGNQVPGEVDFTLDLRSPEESLRLVIDRISVMGEVASQAGVAAAD 3211	RA
RA	-!- FUNCTION: Involved in the anaerobic utilization of allantoin. -!- PATHWAY: Degradation of allantoin (purine catabolism); second step. -!- Converts allantate to (S)-ureidoglycolate and ammonia.	CC	202 ELRLTCGKITAEPNVANVIPGRVQFSIDIRHQHQHVLQFPHDMVALINGICLQKGIRAV 3212	RA
RA	-!- INDUCTION: Expression is very low in excess nitrogen (glutamate plus ammonia) and is induced during limiting-nitrogen conditions (glutamate). Expression is further induced when allantoin is added during limiting-nitrogen conditions.	CC	202 VDEFFNLSPVQLAPTMVDAVREASALQFTHRDIISSGAGHDMSMFIAQVTDGMVFVPSRA 3812	RA
RA	-!- SIMILARITY: Belongs to peptidase family M40.	CC	202 IDEYMRRIEPVPMDERLKAAFAETALENGFSCEEEMVSAGAGHDAQMIGRRYPACMLFVPSRG 3813	RA
RA	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce or send an email to license@isb-sib.ch).	CC	202 GRSHVPEEWTDFDLRKGTTEVVLRVMKAL 410	RA
RA	EMBL; 299120; CAB15243.1; -. PIR; G70017; G70017. DR MEROPS; M40.UNW; --. DR Subtilist; BG13994; pucF. DR InterPro; IPR002933; Peptidase M20. PFam; PF01546; Peptidase M20; $\overline{1}$. DR Hydrolase; Purine metabolism; Complete proteome. KW SEQUENCE; 412 AA; 45519 MW; 852EC0952E4B5802 CRC64;	DR	202 SAEQPGVTRLYTPEHAARETLIAAMRAALSVERDALGN1IGRREGTDPPELPAAVGS 81	RA
RA	Query Match 29.4%; Score 602.5; DB 1; Length 412; Best Local Similarity 35.2%; Pred. No. 8.6e-32; Matches 137; Conservative 70; Mismatches 179; Indels 3; Gaps 2;	Db	202 SAEGPGVTRLYTPEHAARETLIAAMRAALSVERDALGN1IGRREGTDPPELPAAVGS 81	RA
RA	Qy 22 SAEQPGVTRLYTPEHAARETLIAAMRAALSVERDALGN1IGRREGTDPPELPAAVGS 81 Db 25 SDG-GVTRLYTKEWMDAQLAVKTEMSSFGLETRFDDVGNVFGLSGTQSPDEVVTGS 83	DR	202 HFDSVRNGGMFDGTTAGWVCALEAARVMLENGYVNTRPFEFIATVEEGARFSSGMLGGRA 141	RA
RA	Qy 82 HIDTVINGKYDGAYGVLAAMLALKQKLTGYGPKKTLEAVSLCEEEGSRFPMTYWGSQN 143 Db 84 HIDTVINGKYDGAYGVLAAMLALKQKLTGYGPKKTLEAVSLCEEEGSRFPMTYWGSQN 143	DR	202 IAGLVADRELDLSLVDEDGVSVRQAATAFGLKPGELQAAARSAAADRRAFILEHLHQGPILE 201	RA
RA	Qy 142 IAGLVADRELDLSLVDEDGVSVRQAATAFGLKPGELQAAARSAAADRRAFILEHLHQGPILE 201 Db 144 MTGVFSEQDAKEPRDESGVSLQTAQMHESGFGKGVFO SAYRT--DISAFVELHLHQQKITLE 201	DR	202 AAC000157; AAC73618.1; -- EMBL; U82664; AAB40268.1; -- EMBL; U82664; AAB40268.1; -- DR EMBL; U82664; AAB40268.1; --	RA

Query Match	Score 719.5; DB 1; Length 409;
Best Local Similarity	38.6%; Pred. No. 2.6e-39;
Matches 155; Conservative 71; Mismatches 165; Indels 11; Gaps 4;	
7 QAARIEKEIRELRSRFSAE-GPGVTRLTYTPEHAAARETLIAAMKAALSVERDALGNIIG 65	
3 QGERLWQRMLMELGEVGKQPSGGVVTRLSFTAERRAKDLVASYMREAGLFVYEDAAGNLIG 62	
66 RREGTDPELPAAIVGSHFDSVRNGGMFDGTAGVVCALEAARVMLENGYVNRRHPFEFFIAIV 125	
63 RKEGTNPDAVVVLVGSILDSVYNGGCCDFGPLGVLAGDEVVTMNEHGVVTHPIEVVAFT 122	
126 EEEGARFSSGMLGGRAIAGLVADRELDSLWVDEDGVSVRQAATAFGLKPGCELOQAARSAAAD 185	
123 DEEGARFRFGMIGSRAMAGTLPPAECR-DAEGISLAEMKQAGLDPDRILPQAARKPGT 181	
186 LRAFIELDHIEQPILEQQEIQIGVVTSIVGVRALRVAVKGRSDHAGTTPMHLRQDALVPA 245	
182 VKAYVELHIEQGRVLEAGLPLVGVITVGIAGLIWVKFTIAGPAEHAGATPMSLRRDPMAAA 241	
246 ALMV---REVNRFVNNEIADGTVATVGHLLTVAPGGGNQVPGEVDFTLDRSPHEESLRVLL 301	
242 AQIIIVIEEARR----TGTTVGTVGQLHVYPGGINVIPIERVEFVFLDLRLKAEVRDQV 296	
302 IDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMVDAVREAASALQFTHDISSGAGH 361	
297 WKAIAVRAETIAKERNVRLTTERLQEMAPVLCSEVVKQAAERACKQLGYPFWLPSGAAH 356	
362 DSMFIAQVTDVGMVFVPSRAGRSHVPEEWTDFDDLRKGTEVV 403	
357 DGVQLAPICPIGMIFVFRSQDGVSHPSPAEWSTKEDCAVGAEVL 398	
RESULT 4	
Y Y588_HAEIN STANDARD PRT; 411 AA.	
C Q57051; O05027; 01-NOV-1997 (Rel. 35, Created)	
C 01-NOV-1997 (Rel. 35, Last sequence update)	
C 28-FEB-2003 (Rel. 41, Last annotation update)	
NCBI_TaxID=727; Protein H10588.	
SEQUENCE FROM N.A.	
STRAIN=Rd / RW20 / ATCC 51907; MEDLINE=95350630; PubMed=7542800;	
Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., McKenney K., Sutton G., FitzHugh W., Fields C.A., Gocayne J.D., Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M., Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D., Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C., Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M., Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Venter J.C.; "Whole-genome random sequencing and assembly of <i>Haemophilus influenzae</i> Rd."; Science 269:496-512 (1995).	
[2] IDENTIFICATION BY MASS SPECTROMETRY. MEDLINE=20137488; PubMed=10675023; Langen H., Takacs B., Evers S., Berndt P., Lahm H.W., Wipf B., Gray C., Fountoulakis M.; "Two-dimensional map of the proteome of <i>Haemophilus influenzae</i> ."; Electrophoresis 21:411-429(2000). -!- SIMILARITY: Belongs to peptidase family M40.	
SEQUENCE FROM N.A.	
ALLC_BACSU ID_ALLC_BACSU STANDARD; PRT; 412 AA.	
AC 032149; DT 16-OCT-2001 (Rel. 40, Created)	
DB 246 SLAIERAAIQAGHSTVATVGNITAKPGVMNVPGYCELLDIRGTHVQARDSFELLQEE 305	
QY 309 VGEVASQAGVAADVDEFPNLSPVQLAPTMVDAVREAASALQFTHDISSGAGHDSMFIAQ 368	
DB 306 ISKYSEKRGLLIELQLISKDNPILPENMVNQIAETAHSLGYSYEIMPSGAGHDAMHMAT 365	
QY 369 VTDVGMVFVPSRAGRSHVPEEWTDFDDLRKGTEVVLRVM 407	
DB 366 LCPTGMIFIPSHLGGISHNPLEFTDWKDIIEAGIKVLQKV1 404	
RESULT 5	
ALLC_BACSU ID_ALLC_BACSU STANDARD; PRT; 412 AA.	
AC 032149; DT 16-OCT-2001 (Rel. 40, Last sequence update)	
DB 246 SLAIERAAIQAGHSTVATVGNITAKPGVMNVPGYCELLDIRGTHVQARDSFELLQEE 305	
QY 309 VGEVASQAGVAADVDEFPNLSPVQLAPTMVDAVREAASALQFTHDISSGAGHDSMFIAQ 368	
DB 306 ISKYSEKRGLLIELQLISKDNPILPENMVNQIAETAHSLGYSYEIMPSGAGHDAMHMAT 365	
QY 369 VTDVGMVFVPSRAGRSHVPEEWTDFDDLRKGTEVVLRVM 407	
DB 366 LCPTGMIFIPSHLGGISHNPLEFTDWKDIIEAGIKVLQKV1 404	
SEQUENCE FROM N.A.	
RC STRAIN=168; RX MEDLINE=98044033; PubMed=9384377;	
RA Kunst F., Ogabawara N.,' Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bebbieres P., Bolotin A., Borchert S., Borriis R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoff A., Ehrlich S.D., Emmerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Galizzi A., Galleron N., Fritz C., Fujita M., Fujita Y., Fuma S., Goffeau A., Golightly E.J., Grandi G., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,	
This SWISS-PROT entry is copyright. It is produced through a collaboration	

QY	126 EEEGARFSSGMLGGRAIAGLVADREILDSLVDDEDGVSVRQAATAFGLKPKGELOAAARSAAAD
Db	127 : : : : : : : : : : : : : : :
QY	128 AAARSAAADLRAFIELHIEQCPILEQQIEIGVVTIVGVRAVKGRSDDHAGTTPMHL 237
Db	129 QSIREFIDIKHYFEMHIEQGPYLEKQNNYPIGIVSGTAGPSWFKVRLVGAEAGHACTVPMSL 239
QY	130 RQDALVPAALM/REVNRF-VNEIADGTVATVGHLTAVAPGGGNQVPGEVDFTLDLRSPHEE 296
Db	131 240 RKDPLVGRAAEVKEVETLCMNDPNAPTVGTGRIAFAFPGGSMNIIPESVEFTLDIRDIELE 299
QY	132 297 SLRVLDLDRISVMVGEVASQAGVVAADVDEFNLSPVQLAPTMVDAVREAASALQFTTHRDIS 356
Db	133 300 RRNKIEKIEEKIKLVLNSNTRGLEQIEKNMAAVPVKCSENLINSLKQSCKELEIDAPIIV 359
QY	134 357 SGAGHDSSMFLIAQVTDVGGMVFVPSRAGRSHVPEEWTDFDLRKGTTEVV 403
Db	135 360 SGAGHDAMFLAETIEIGMVFVRCRNGISHSPKEWAIEIDDILTRGKYL 406
QY	136 357 DSVQLAPICPIGMIFVRSQDGVSHSPAESTKEDCAAGAEVL 398
RESULT 2	
AMB2_BACST	AMB2_BACST STANDARD; PRT; 409 AA.
Q53389;	01-NOV-1997 (Rel. 35, Created)
	01-NOV-1997 (Rel. 35, Last sequence update)
	28-FEB-2003 (Rel. 41, Last annotation update)
DE	N-carbamoyl-L-amino acid hydrolase (EC 3.5.1.87) (L-carbamoylase).
EN	AMAB.
DS	Bacillus stearothermophilus.
DC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.
DX	[1]
RN	NCBI_TaxID=1422;
RP	NCBI_TaxID=1422;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	RP SEQUENCE OF 60-409 FROM N.A.
XX	STRAIN=NCIB 8224;
RA	MEDLINE=94115050; PubMed=7764340;
RT	Mukohara Y., Ishikawa T., Watabe K., Nakamura H.;
RT	"Molecular cloning and sequencing of the gene for a thermostable N-
RT	carbamoyl-L-amino acid amidohydrolase from <i>Bacillus stearothermophilus</i>
RT	strain NS1122A."
RT	BioTechnol. Biotechnol. 57:1935-1937(1993).
RT	-!- CATALYTIC ACTIVITY: N-carbamoyl-L-2-amino acid (a 2-ureido
RT	carboxylate) + H(2)O = L-2-amino acid + NH(3) + CO(2).
RT	-!- COFACTOR: Manganese or nickel or cobalt.
RT	-!- SIMILARITY: Belongs to peptidase family M40.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	EMBL; AAC60456.1; -.
CC	PIR; JN0885; JN0885.
CC	MEROPS; M40_UNW; -.
CC	InterPro; IPR002933; Peptidase_M20.
CC	Pfam; PF01546; Peptidase_M20; -.
CC	Hydrolyase.
CC	SEQUENCE 409 AA; 44248 MW; 332515C62A1A2F31 CRC64;
CC	Query Match 35.5%; Score 726.5; DB 1; Length 409;
CC	Best Local Similarity 38.8%; Pred. No. 9.2e-40;
CC	Matches 156; Conservative 72; Mismatches 163; Indels 11; Gaps 4;
CC	Sequence 3Q
CC	7 QAAIEKEIRELSRSFSAE-GPGVTRLTYTPEHAARETLIAAMKAALSVERDALGNIIG 65
CC	3 QGERLWQRLMELGEVKQPSGGVTRLSFTAERRAKDLVASYMREAGLFVYEDAAGNLIG 62
CC	66 RREGTDPPELPAIAVGSHFDSDVRYNGGMFDGTAGVVALEAARVMLENGYVNRRHFFIFIIV 125
CC	63 RKEGTNPDATAVVLVGSNLDVSYNGGCFCGPLGVLAGVEVVTMNEHGVVTHHPIEVVAFT 122
DR	DR EMBL; X74289; CAA52341.1; -.
DR	DR MEROPS; M40_UNW; -.
DR	DR InterPro; IPR002933; Peptidase_M20; -.
DR	DR Pfam; PF01546; Peptidase_M20; -.
KW	KW Hydrolase.
FT	FT CONFLICT 60 62 LIG -> DRE (IN REF. 2).
SQ	SEQUENCE 409 AA; 44166 MW; 9A91F767C9AEFBBC CRC64;

This Page Blank (up to)

Job time : 22 secs

RESULT 15

AB3457
N-carbamoyl-L-amino acid amidohydrolase (EC 3.5.1.-) [imported] - Brucella melitensis (s
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 15-Feb-2002
C;Accession: AB3457
R;DelVecchio, V.G.; Kapatrai, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AB3457
A;Status: preliminary
A;Map position: I
A;Cross-references: GB:AE008917; PIDN:AAL52824.1; PID:g17983664; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI1643
A;Molecule type: DNA
A;Residues: 1-415 <KUR>
C;Superfamily: N-carbamyl-L-amino acid amidohydrolase
C;Keywords: hydrolase

Query Match 26.2%; Score 536.5; DB 2; Length 415;
Best Local Similarity 32.1%; Pred. No. 1.2e-28;
Matches 132; Conservative 73; Mismatches 177; Indels 29; Gaps 8;

Qy 10 RIEKEIRELRSRFAEGPGV- ---TRLTYTPEHAAARETLIAAMKAALSVERDALGNIIIG 65
Db 13 RLWDSSMEMAKI- --GPGLRGGNRQTLTDEDGEGRRLFQNWC EKGLSMGVDTMGNMFF 69

Qy 66 RREGTDPPELPATAVGSHFDSVRNGGMFDGTAGVVCALEAARVMLENGYVNRRHPPEFLAIV 125
Db 70 TRPGEEDSDADPVYMGSHLDTQPTGGKF DGVLGVLGLEVMRRTLNDMNIRT KRP IVVVWNT 129

Qy 126 EEEGARFSSGMLGGRAIAGLVADRELDLSVDEDGVSVRQAATAFGLKPGELOAAARS AAD 185
Db 130 NEEGTRFAPAMLA SGVFGVLDQN WAYERTDAGKTKFGEELVRIGWK-GDEPVGSRK--- 185

Qy 186 LRAFILELHIEQGPILLEQEIQIEGVUTSIVGVRALRVAKGRSDHAGTTPMHLRQDALVPA 245
Db 186 IHAMFELHIEQGPILAEHKDIDGVVTHGOGLWWLQVTLTGKEAHTGSTPMRMKRN---A 241

Qy 246 ALMVR ENVRFVNEIA ---DGTVA TUVGHLTVA PGGGNQVPGEVDFTDLRSPHEESLRV L 301
Db 242 SLGLGKLLQLVN ETIAMAHOPDAVGGVGHIDVSPNSRNVLPGQIVFTVDFRSPNQATL DGM 301

Qy 302 IDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMVDAVREAAASALQFTHRDISSGAGH 361
Db 302 KARFEKEAPKIAEELGIGIEEVAGHFDPVTFTDGCVEAIRNAERLGYSRNI VSGAGH 361

Qy 362 DSMFIAQVTDVG MVFVPSRA GRSH----VPEEWTFDDLRKGTEVVL RV 407
Db 362 DACWVNRVAPTA MCMPCV DGL SHNED DISKEWA S----AGTDVLLH A V 407

A; Molecule type: DNA A; Residues: 1-415 <KUR> A; Cross-references: GB:AE007869; PIDN:AAK80122.1; PID:g15157556; GSPDB:GN00169 C; Genetics: A; Gene: AGR C 4327 A; Map position: circular chromosome C; Superfamily: N-carbamyl-L-amino acid amidohydrolase	Qy 59 ALIGNIIGREGTDPPELPAIAVGSHFDSDVRNGGMFDGTAGVVCALEAARVMLENGYVNRRHP 118 Db 61 AVGNLRSRLESRAPGARTLYIGSHLDTVNAGRYDGIIVVFGYALVEALRE---RELP 116
Query Match 27.4%; Score 562; DB 2; Length 415; Best Local Similarity 35.2%; Pred. No. 2.3e-30; Matches 140; Conservative 59; Mismatches 169; Indels 30; Gaps 9;	Qy 119 F--EFFIAIEEEGARFSSGMLGGRAAGLVADRELDLSVDEDGVSVRQAATAFGLKPGEL 176 Db 117 FHLEVLFSEEEGVYGVSFIGSRLAVG-TAD-ELLTVDRENQNTVRDAIVGYGLNPDEL 174
Qy 177 QAAARSAAADLRAFIELHIEQGPILEQEQQIEGVVTSIVGVRAVKGRSRHAGTTPMH 236 Db 175 FGAGAEDRPL-GYLEIHIIEQGPVLQDQGAAVGVSAIVQGSRLTLHFTGRASHAGTTPMH 233	Qy 177 QAAARSAAADLRAFIELHIEQGPILEQEQQIEGVVTSIVGVRAVKGRSRHAGTTPMH 236 Db 175 FGAGAEDRPL-GYLEIHIIEQGPVLQDQGAAVGVSAIVQGSRLTLHFTGRASHAGTTPMH 233
Qy 237 1RQDALVPAAALMVRENVNRFVNEIADGTVATVGHLTVAFFGGNNQVPGEVDETLDLRSPHEE 296 Db 234 IRRDALAAAARFIVGAEDLANR-TPGLVATVGMIEAKPGAGNVIAEGVSCLDIRHADDA 292	Qy 237 1RQDALVPAAALMVRENVNRFVNEIADGTVATVGHLTVAFFGGNNQVPGEVDETLDLRSPHEE 296 Db 234 IRRDALAAAARFIVGAEDLANR-TPGLVATVGMIEAKPGAGNVIAEGVSCLDIRHADDA 292
Qy 297 SLRVLIDRISVMVGEVASQAGVAADVDDEFNLSPVQLAPTMVDAVREAASALQFTHRDIS 356 Db 293 VRAQSLQELLALAAREAAARQVACSVTPRMAEAAPVMPAPSILRALLHQAAEQGLTHPELV 352	Qy 297 SLRVLIDRISVMVGEVASQAGVAADVDDEFNLSPVQLAPTMVDAVREAASALQFTHRDIS 356 Db 293 VRAQSLQELLALAAREAAARQVACSVTPRMAEAAPVMPAPSILRALLHQAAEQGLTHPELV 352
Qy 357 SGAGHDMSMFIAQVTDVGMVFVPSRAGRSHVPEEWTDFDDLRKGTEVVLRVMKAL 410 Db 353 SGAGHDQDAQIMAQRMPPAAMLFLRSPPNALSHHPDEMAEPGDVAAGLVRGTRPLELL 406	Qy 357 SGAGHDMSMFIAQVTDVGMVFVPSRAGRSHVPEEWTDFDDLRKGTEVVLRVMKAL 410 Db 353 SGAGHDQDAQIMAQRMPPAAMLFLRSPPNALSHHPDEMAEPGDVAAGLVRGTRPLELL 406
Qy RESULT 14 Db H95999 probable N-carbamyl-L-amino acid amidohydrolase protein (EC 3.5.1.-) [imported] - Sinorhizobium meliloti	Qy RESULT 14 Db H95999 probable N-carbamyl-L-amino acid amidohydrolase protein (EC 3.5.1.-) [imported] - Sinorhizobium meliloti C; Species: Sinorhizobium meliloti C; Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001 C; Accession: H95999 R; Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T. Science 293, 668-672, 2001 A; Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, R.; Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K. A; Title: The composite genome of the legume symbiont Sinorhizobium meliloti. A; Reference number: A96039; PMID:21368234; PMID:11474104 A; Status: preliminary A; Molecule type: DNA A; Residues: 1-414 <KUR> A; Cross-references: GB:AL591985; PIDN:CAC49664.1; PID:g15141151; GSPDB:GN00167 A; Experimental source: strain 1021, megaplasmid pSymB R; Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.
Qy 258 EIADGTVATVGHLTVAFFGGNNQVPGEVDEFTDLRSPHE---ESLRVLIDRISVMVGEVAS 314 Db 258 AHQPGAVAGVGQMIFTPSNRNTVLPGVVFTIDLRTPSQAKLDMRAIFER---EVPVIAE 314	Qy 258 EIADGTVATVGHLTVAFFGGNNQVPGEVDEFTDLRSPHE---ESLRVLIDRISVMVGEVAS 314 Db 258 AHQPGAVAGVGQMIFTPSNRNTVLPGVVFTIDLRTPSQAKLDMRAIFER---EVPVIAE 314
Qy 315 QAGVAADVDEFFNLSPVQLAPTMVDAVREAASALQFTHRDISSGAGHDSMFMIAQVTDVGM 374 Db 315 ELGVGCSIEAIGHFDPVTFDPVILVGRVRSAAERLGTYMDIIISGAGHDACWTARVAPSTM 374	Qy 315 QAGVAADVDEFFNLSPVQLAPTMVDAVREAASALQFTHRDISSGAGHDSMFMIAQVTDVGM 374 Db 315 ELGVGCSIEAIGHFDPVTFDPVILVGRVRSAAERLGTYMDIIISGAGHDACWTARVAPSTM 374
Qy 375 VFVPSRAGRSH---VPEEWTDFDDLRKGTEVVLRVMM 407 Db 375 IFCPGVGLSHNEEEEISPEWA----AAGCDVLLHAV 407	Qy 375 VFVPSRAGRSH---VPEEWTDFDDLRKGTEVVLRVMM 407 Db 375 IFCPGVGLSHNEEEEISPEWA----AAGCDVLLHAV 407
Qy F75429 N-carbamyl-L-amino acid amidohydrolase (EC 3.5.1.-) DR1154 [similarity] - Deinococcus radiodurans C; Species: Deinococcus radiodurans C; Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 08-Sep-2000 C; Accession: F75429 R; White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M.; Smith, H.O.; Venter, J.C.; Fraser, C.M. Science 286, 1571-1577, 1999 A; Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1. A; Reference number: A75250; MUID:20036896; PMID:10567266 A; Accession: F75429 A; Status: preliminary A; Molecule type: DNA A; Residues: 1-416 <WHI> A; Cross-references: GB:AE001965; PIDN:AAF10728.1; PID:g6458892; NID:g6458892; GSPDB:GN00169 C; Genetics: A; Gene: DR1154 A; Map position: 1 C; Superfamily: N-carbamyl-L-amino acid amidohydrolase C; Keywords: hydrolase	Qy 27 GVTRLTYPEHAAARETLLIAAMKAALSVREDALGNIIGREGTDPPELPAIAVGSHFDSSV 86 Db 31 GLARLTLLDDDRMRDLFVNWNCRDAGLCVIEDLGNIFGRREGDPSLPPVLGSHLDTQ 90
Qy 87 RNNGMFDTAGVVCALEAARVMLENGYVNRRHPFEFIAIVEEGARFSSGMLGGRAIAGLV 146 Db 91 INGGRFDDGIAVLSGLEVIRTLNDVGHVTRPPIVWNNEEGARFSPPMVASGCFVGAY 150	Qy 87 RNNGMFDTAGVVCALEAARVMLENGYVNRRHPFEFIAIVEEGARFSSGMLGGRAIAGLV 146 Db 91 INGGRFDDGIAVLSGLEVIRTLNDVGHVTRPPIVWNNEEGARFSPPMVASGCFVGAY 150
Qy 147 ADRELDSSLVDEDGVSVRQAATAFGLKPGELQAAARS-----AADLRAFIELHIEQGPIL 200 Db 151 DGDWVKELVSDDG----ARFG---AEELERIGYNGNKPORAGEIDAYYELHIEQGPIL 200	Qy 147 ADRELDSSLVDEDGVSVRQAATAFGLKPGELQAAARS-----AADLRAFIELHIEQGPIL 200 Db 151 DGDWVKELVSDDG----ARFG---AEELERIGYNGNKPORAGEIDAYYELHIEQGPIL 200
Qy 201 EQEQIEIGGVVTSIVGGVRAVKGRSDDHAAGTTPMHLRQDALVPAALMREVNRFVNEIA 260	Qy 201 EQEQIEIGGVVTSIVGGVRAVKGRSDDHAAGTTPMHLRQDALVPAALMREVNRFVNEIA 260

Qy 326 FNLSPVQLAPTMVDAVREAA SALQFTHRD ISSGAGHD SSMFIAQVTDGV MVFVPSRAGRSH 385
 Db 327 QALAESPCDPSLMLGDLDESLADLGLPARRLPSGAGHDAMVMAALCPTAMLFIRCEGGISH 386

RESULT 11
 N-carbamoyl-beta-alanine amidohydrolase [imported] - Agrobacterium tumefaciens (strain AG2869)

C;Species: Agrobacterium tumefaciens
 C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell- i; Karp, P.; Romero, P.; Zhang, S.

C;Accession: AG2869
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A;Reference number: AB2577; MUID:21608550; PMID:11743193
 A;Accession: AG2869
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.

C;Species: Bacillus subtilis
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 08-Sep-2000

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chc A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleriech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullio, M.F. Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauel, Y.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleicher, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serorakechi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Togano, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yata, K.; Yoshida, K. A;Cross-references: GB:Z99120; GB:AL009126; NID:92635613; PID:92635750
 A;Experimental source: strain 168
 A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. A;Reference number: A69580; MUID:98044033; PMID:9384377
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-412 <RUN>
 A;Cross-references: GB:Z99120; GB:AL009126; NID:92635613; PID:92635750
 A;Experimental source: strain 168
 C;Genetics:
 A;Gene: YurH
 C;Superfamily: N-carbamoyl-L-amino acid amidohydrolase
 C;Keywords: hydrolase
 C;Keywords: hydrolase

Query Match 29.4%; Score 602.5; DB 1; Length 412;
 Best Local Similarity 35.2%; Pred. No. 4.4e-33;
 Matches 137; Conservative 70; Mismatches 179; Indels 3; Gaps 2;

Db 22 SAEGPGCVTRLTYPEHAAARETLLIAAMKAAALSVERDALGNITGRREGTDPELPALAVGS 81
 Db 25 SADG-GYTRLLTYKEWMDAQLAVKTEMSSFGLETRFDVGNVFGRLSGTQSPDEVITRGS 83

Qy 82 HFDSVRNGGMFDGTAGVVCALEAARVMLENGYVNRPFFEIAIVEEEGARFSSGMLGGRA 141
 Db 84 HIDTVINGKYDGA YGVLAAMLLALKQLKETYGA PKTLEAVS LCEE EEGSRFPMTYWGSGN 143

Qy 142 IAGLVADRELDLVLVEDDGVSYRQAATAFGLKPGELOQAARSAA DLRAFIELHIEQGPILE 201
 Db 144 MTGVSEQDAKEPRDES GVSLOQTAMHESFGKGKVFO SAYRT--DISAFVELHIEQKTLE 201

Qy 202 QE QIEGIVVTSIVGVRLRVAKGRSDHAGTTPMHLRQD ALVEA ALMVR EVNRFVNBIAD 261
 Db 202 MSGRD LGIVVTSIAGQRRYLVTL EGC NHAGTTSMKWRKDPLAASSRITHELLRSDELPD 261

RESULT 12
 A97646
 n-carbamoyl-beta-alanine amidohydrolase (PA0444) [imported] - Agrobacterium tumefaciens

C;Species: Agrobacterium tumefaciens
 C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002

R;Goodner, B.; Hinkle, G.; Gartung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; IDEYMRIEPVPMDERILKA AAFETALENGFSCEEMVSGAGHDAQMIGRRYPACMLFVPSRG 381

Qy 382 GRSHVPEEWTFDDLRKGTEVVL RVMKAL 410
 Db 382 GVSHSPKEYTSARQLEIGVRA LDTLYKL 410

A;Accession: A97646
 A;Reference number: A97359; MUID:21608551; PMID:11743194
 A;Status: preliminary

11. M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, N; Nature 413, 523-527, 2001
 A; Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A; Reference number: AB0001; MUID:21470413; PMID:11586360

A; Accession: AH0394
 A; Status: preliminary

A; Molecule type: DNA

A; Cross-references: GB:AL590842; PIDN: CAC92484.1; PID: 915981185; GSPDB: GN00175

C; Genetics:
 A; Gene: YPO3249

C; Superfamily: N-carbamyl-L-amino acid amidohydrolase

Query Match Score 642.5; DB 2; Length 430;
 Best Local Similarity 36.6%; Pred. No. 9.7e-36;
 Matches 145; Conservative 66; Mismatches 172; Indels 13; Gaps 3;

Qy 18 LSREFSAEGPGVTRLTYTPEHAAARETLIAAMKAALSVREDALGNITGRREGTDPELPAI 77
 Db 24 LAISESPEGLTRYVLSPEHLRANRQVGEMQAVGMQVWQDFTVGNICGRYEGRQPDAAI 83

Qy 78 AVGSHFDSSVRNGGMFDGTAGVVALEAARVMLENQYVNRRH-----PFEFIAVEEEGAR 131
 Db 84 LLGSHLDTVRNAGRHYDGMLGVLTALEVV-----GYLHRHQQLRPLPVATEVIGFADEEGTR 137

Qy 132 FSSGMLGGRAIAGLVADRELDSLVDDEGVSVRQAATAFGLKPGELQAAARSAAADLRAFIE 191
 Db 138 FGITLLGSKGYTGRWPVEWLN-TDADGISVAQMVRAQMLDPMDIGQSARAANAFCAYLE 196

RESULT 9
 H87571

N-carbamyl-L-amino acid amidohydrolase [imported] - Caulobacter crescentus

C; Species: Caulobacter crescentus

C; Date: 20-Apr-2001 #text_change 10-May-2001

C; Accession: H87571

R; Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonin, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A; Title: Complete Genome Sequence of Caulobacter crescentus.

A; Reference number: A87249; PMID:21173698; PMID:11259647

A; Accession: H87571

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-427 <STO>

C; Genetics:

A; Gene: CC2603

C; Superfamily: N-carbamyl-L-amino acid amidohydrolase

Query Match Score 612.5; DB 2; Length 427;

Best Local Similarity 36.1%; Pred. No. 9.9e-34;

Matches 139; Conservative 70; Mismatches 171; Indels 5; Gaps 3;

Qy 252 VNRFVNNEIADGTATVGHLTVAAPGGGNQVPGEVDFTLDLRSPEESLRVLIDRISVMVGE 311
 Db 257 VEALTAQQGEHLVATVGTLTCLPGAVNVIPQVRLTDIRGPNDRGVNDLTRLAAEAA 316

Qy 312 VASQAGVAADVDFFNLSPVQQLAPTMVDAREAAASALQFTHRDIISSGAGHDMSMFIAQWTD 371
 Db 317 IATRRGIFTAAEGFYRIKATACDSALQQCISIQVQRCLALPPSGAGHDAAIMAECWP 376

Qy 372 VGMVFVPSRAGRSHVPEEWTDFDDLRKGTEVVVLRVM 407
 Db 377 VGMFLFRCKGGVSHHPDESVTSSDVAVIAQAYLEAV 412

RESULT 8
 C83591
 N-carbamoyl-beta-alanine amidohydrolase PA0444 [imported] - Pseudomonas aeruginosa (strain PA01)
 C; Species: Pseudomonas aeruginosa
 C; Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C; Accession: C83591
 R; Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Coulter, S.N.; Folger, K.R.; Warrener, P.; Hickey, M.J.; B.Z. adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, B.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A; Title: Complete Genome Sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
 A; Reference number: A82950; MUID:20437337; PMID:10984043
 A; Accession: C83591
 A; Status: preliminary
 A; Molecule type: DNA
 A; Cross-references: GB:AE004481; GB:AE004091; NID:99946293; PIDN:ARG03833.1; GSPDB:GN001
 A; Experimental source: strain PA01
 C; Genetics:
 A; Gene: PA0444
 C; Superfamily: N-carbamyl-L-amino acid amidohydrolase

Query Match Score 639.5; DB 2; Length 427;
 Best Local Similarity 38.1%; Pred. No. 1.5e-35;
 Matches 156; Conservative 54; Mismatches 190; Indels 9; Gaps 3;

Qy 2 TLQKAQARIEKEIRELSRFSAEQP-GVTRLTYTPEHAAARETLIAAMKAALSVERDAL 60

Db 11 TQRHIDGQRLWQSLMDLARLGATAKGGVCRLLSDLDRQARDLFLQWCEAAGCTVSVDRV 70
 Qy 61 GNIIGRREGTDPELPAIAGSHFDSSVRNGGMFDGTAGVVALEAARVMLENGYVNRRHPFE 120
 Db 71 GNIFARRPGRNPDLPPMTGSHIDTQPTGGCFGVNAGLEVIRTLNDLGVETEAPLE 130
 Qy 121 FIAIVEEGARFSSGMLGGRAIAGLVADRELDSLVDDEGVSVRQAATAFGLKPGELQAAA 180
 Db 131 VVWTNEEGSRFAPCMGSGVFGKFTLEETLAKRDAGSVGEALDAIG-----YAGA 184
 Qy 181 RSAA--DLRAFIELDHIEQGPLQEQQIEIGVVTTSIVGVRALRVAKGRSDHAGTTPMHHLR 238
 Db 185 RDCLGHGPVGAYFEAHIEQGPLDEEEKTIGVVLGALGQKWFDSLRLRGVEAHAGGTPMHLR 244
 Qy 239 QDALVPAALMVRREVNRNFVNEIADGTATVGHLTVAAPGGGNQVPGEVDFTLDLRSPEESL 298
 Db 245 KDALVGAAVAVEAVNRAALGHOPHACGTVGCLHAYPGSRNVPGEVKMTLDFRHQLOPERL 304
 Qy 299 RVLIDRISVMVGEVASQAGVAADVDDEFNLSPVQQLAPTMVDAREAAASALQFTHRDIISSG 358
 Db 305 DSMIAEVROQVIAATCEKHGLQYELVPTADFPFLYFDQCCVGAVERAAQALGMPQMDIVSG 364
 Qy 359 AGHDSMFIAQVTDVGMVFVPSRAGRSHVPEEWTDFDDLRKGTEVVVLRVM 407
 Db 365 AGHDAIFLAEELGPAGMIFVPCENGISHNEIEENASPDDLAAGCAVLRAM 413

RESULT 9
 H87571
 N-carbamyl-L-amino acid amidohydrolase [imported] - Caulobacter crescentus
 C; Species: Caulobacter crescentus
 C; Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C; Accession: H87571
 R; Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonin, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A; Title: Complete Genome Sequence of Caulobacter crescentus.
 A; Reference number: A87249; PMID:21173698; PMID:11259647
 A; Accession: H87571
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-427 <STO>
 C; Genetics:
 A; Gene: CC2603
 C; Superfamily: N-carbamyl-L-amino acid amidohydrolase

Query Match Score 612.5; DB 2; Length 427;
 Best Local Similarity 36.1%; Pred. No. 9.9e-34;
 Matches 139; Conservative 70; Mismatches 171; Indels 5; Gaps 3;

Qy 28 VTRLTYTPEHAAARETLLIAAMKAALSVERDALGNILIGRREGTDPELPAIAGVSHFDSSVR 87
 Db 30 LTRRFLTAHGAALDLAGWMAEAGMSARRDAANLIGRYEGETHAKALITIGSHIDSVR 89

Qy 88 NGGMFDGTAGVVALEAARVMLENGYVNRPF--EFIAIVEEGARFSSGMLGGRAIAGL 145
 Db 90 NGGRYDGPLGIMLGIDVVEALHRAAG--RRLPFAIEVVAFGDEEGSRFPASMSCSRAIAGT 147

Qy 146 VADRELDSSLVDDEGVSVRQAATAFGLKPGELQAAAARDLRAFIELDHIEQGPLQEQQI 205
 Db 148 LDATALE-MKDAEGSVVAEALAAFFGGDPANTIASAARRPEEVLAFLAEHIEQGPVLEAEGL 206

Qy 206 EIGVVTTSIVGVRALRVAKGRSDHAGTTPMHLRQDALVPAALMVRVNFEVNEIADGTVA 265
 Db 207 ALGVVTTAAQKRLMVRITGMAGHGTTPMALRKDPGEAAEAIALALERICRAGTDGLVGC 266

Qy 266 TVGHLTVAPGGGNQVPGEVDFTLDLRSPEESLRLVIDRISVMVGEVASQAGVAADVDEF 325
 Db 267 TVGRMTALPGAFNVIPGAIEFSMDIRAEFTAATDAVEAITAEIHAAARDLSATVTLML 326

Db	121	EVVAFCEEGSRTNDGLFGSRGMVKVPEDLQK-VDDNNVTRYEALKTFRGFGIDPFDTH	179	Db	242	AQIIIVIEEARR----TGTRVGTVGQLHVPGGINVIPERVEFLDLRDLKAEVRDQV	296		
Qy	178	AAARSAADLRAFIELHIEQPILEQEIQIEGVVTTSIVGVRAVKGRSRSDHAGTTPMHL	237	Qy	302	IDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDVREAAASALQFTHDISSGAGH	361		
Db	180	QSTIREIGDIKHYFEMHIEQGPYLEKNNYPIGIVSGIAGPSWFKVRLVGEAIGHAGTVPMSL	239	Db	297	WKAIAVRAETIAKERNRVTTTERLQEMPPVLCSDEVKRAEAACQKLGYPSFWLPSGAHH	356		
Qy	238	RQDALVPAALMVRREVNRF-VNEIADGTATVGHLTVAAPGGGNQVPGEVDETLDLRSPEEE	296	Qy	362	DSMFTAQYTDVGMVVFVPSRAGRSHVPEEWTDFDDLRLKGTEVV	403		
Db	240	RKDPLVGAEEVIKEVETLCMNDPNAFTVGRRIAAPPGGSNLIIPESVERFLDIRDIELE	299	Db	357	DSVQLAPICPIGMIFVRSQDGVSHSPAEWSTKEDCAGAEVL	398		
Qy	297	SLRVLDRIDSVNVGEVASQAGVAADVDEFFNLSPVQLAPTMVDVREAAASALQFTHDIS	356	RESULT 6					
Db	300	RRNKIEKIEEKIKLKVSNTRGLEQIEKRNMAAVPVKCSENINSLKQSCKELEIDAPIV	359	Db	64079	probable N-carbamyl-L-amino acid amidohydrolase (EC 3.5.1.-) HI0588 [similarity] - Haemophilus influenzae			
Qy	357	SGAGHDSMFIAQWTDGMVFVPSRAGRSHVPEEWTDFDDLRLKGTEVV	403	C;Species:	Haemophilus influenzae				
Db	360	SGAGHDAFLAETITEIGMVVFVRCRNGISHSPKEWAIDDILITGTKVL	406	C;Date:	10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 08-Sep-2000				
Qy	360	SGAGHDSMFIAQWTDGMVFVPSRAGRSHVPEEWTDFDDLRLKGTEVV	403	C;Accession:	D64079				
Db	360	SGAGHDAFLAETITEIGMVVFVRCRNGISHSPKEWAIDDILITGTKVL	406	R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, R.; Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, R.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.					
RESULT 5									
JN0885		N-carbamyl-L-amino acid amidohydrolase (EC 3.5.1.-) [validated] - Bacillus stearothermophilus		A;Title:	Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.				
C;Species:	Bacillus stearothermophilus	C;Date:	10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Sep-2000	A;Reference number:	MUID:95350630; PMID:7542800				
C;Accession:	JN0885; PN0679; I40357	A;Status:	nucleic acid sequence not shown; translation not shown	A;Accession:	D64079				
A;Molecule type:	DNA	A;Molecule type: DNA		A;Residues:	1-411 <TIGR>				
A;Cross-references:	GB:U32740; GB:L42023; NID:91573572; PIDN: AAC22245.1; PID:91573578;	A;Cross-references:	GB:U32740; GB:L42023; NID:91573572; PIDN: AAC22245.1; PID:91573578;	C;Superfamily:	N-carbamyl-L-amino acid amidohydrolase				
A;Accession:	JN0885	C;Keywords:	hydrolyase						
A;Molecule type:	DNA			Query Match	32.7%	Score 671; DB 1;	Length 411;		
A;Residues:	1-409 <MUK>			Best Local Similarity	36.1%	Pred. No. 1.1e-37;			
A;Cross-references:	GB:S67784; NID:9460894; PIDN: AAC60456.1; PID:9460895			Matches	144;	Conservative 80; Mismatches 173; Indels 2; Gaps 2;			
A;Accession:	PN0679			10 RIEKEIRELSRFSAAEGPVTRLTYTPEHAARETLIAAMKAALSVERDALGNIIGRREG	69				
A;Molecule type: protein				7 RVQNLLIEKLAFAISSVNPNELTRLAFTEDDEKAHNMILECLKEYDLISRDSIGNLFIRKAG	66				
A;Residues:	1-13 <MU2>			70 TDPelpaIavGSHFDsvrNGMFDGTAGVVCALEAARVMLENGYVNRPFFBIAVEEG	129				
R;Sakanyan, V.; Desmarez, L.; Legrain, C.; Charlier, D.; Mett, T.; Kochikyan, A.; Savchenko, V.; Ishikawa, T.; Watabe, K.; Nakamura, H.				67 KEDFLPWAFAFGSHIDTVVNAGKFDPGLGSVAGLEILLQLCEQNIQTRYPLELILFTCEES	126				
A;Title:	Gene cloning, sequence analysis, purification, and characterization of a thermo			130 ARFSSGMILGGRATAIGLVADRELLSLVDEDGVSVRQAATAFLGLKPGELOAAARSAAIDLRAF	189				
A;Reference number:	140357; MUID:94113715; PMID:8285691			127 SRFNFATLGSKVCMCGIVNQEKLSSLRDKQKGGLSEAMAEVGMNFMNVNQAKRDAKEFKCF	186				
A;Accession:	I40357			190 IELHIEQQPTILEOEQIEIGVVTTSIVG-VRALRVAVKGRSDHAGTTPMHLRQDALVPAALM	248				
A;Status:	preliminary; translated from GB/EMBL/DDBJ			187 FELHIEQQPRLNEGKTIGVVTGIAAPIRAI-VKIKGQADHSGATAMHYRHDAALLGGSEL	245				
A;Molecule type:	DNA			249 VREVNRFVNEIADGTVATVGHILTAPGGGNQVPGEVDFTLDLRSPEEESLRVLIIDRISVM	308				
A;Residues:	'DRB', '63-198, 'A', '200-219, 'A', '221, 'P', '223-314, 'L', '316-323, 'A', '325-329, 'EV', '33			246 SLAIERAATIQAQGHSTVATVGNNTAKPGVMNNVPGYCELLVDIRGTHVQARDSVPELLQEE	305				
A;Cross-references:	EMBL:X74289; NID:9436795; PID:9436796			309 VGEVASQAGVAADVDEFNLSPVQLAPTMVDVREAAASALQFTHDISSGAGHDMSMFIAQ	368				
C;Description:	catalyzes the hydrolysis of an N-carbamoyl-L-amino acid to ammonia, carbonyl group			306 ISKVSEKRGLLIELQLISKDNPIILPENMVNQIAETAHSLGYSYEIMPSGAGHDAMHMAT	365				
C;Superfamily:	N-carbamyl-L-amino acid amidohydrolase			369 VTDVGMVVFVPSRAGRSHVPEEWTDFDDLRLKGTEVVLRM	407				
C;Keywords:	hydrolyase			366 LCPTGMIFIPSGLGSHNPLEFTDWKLDIEAGIKVLOQVI	404				
Query Match	35.5%	Score 726.5; DB 1; Length 409;		RESULT 7					
Best Local Similarity	38.8%	Pred. No. 2.1e-41; Mismatches 72; Indels 11; Gaps 4;		AH0394					
Matches	156;	Conservative		C;Species:	Yersinia pestis (strain CO92)				
Qy	7	QAARIEKEIRELSRFSAE-GPGVTRLTYTPEHAARETLIAAMKAALSVERDALGNIIG	65	C;Date:	02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001				
Db	3	QGERLWQRLMELGEVKGQPSGGVTRLSETAERRAKDVLASYMREAGLFVYEDAAGNLIIG	62	C;Accession:	AH0394				
Qy	66	RREGTDPelpaIavGSHFDsvrNGMFDGTAGVVCALEAARVMLENGYVNRPFFBIAIV	125	R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.					
Db	63	RKEGTNPDATAVVLVGSHLDSYVNGGCCFDGPLGVLAGVEVQTMNEHGVVTHPIEVWAFT	122	C;Cross-references:	ALMV---REVNRFVNEIADGTVATVGHILTAPGGGNQVPGEVDFTLDLRSPEEESLRV	301			
Qy	126	EEEGARFSSGMLGGRATAIGLVADRELLSLVDEDGVSVRQAATAFLGLKPGELOAAARSAA	185	C;Keywords:	Chillingworth, A.M.; Deno-Tarraga, A.M.; Davies, R.M.; Dougan, G.;				
Db	123	DEEGARFRFGMIGSRAMAGTLPPEALECR-DAEGISLAEAMKOAGLDPDRLPQAARKPGT	181						
Qy	186	LRAFIELHIEQPILEQEIQIEGVVTTSIVGVRAVKGRSRSDHAGTTPMHLRQDALVPA	245						
Db	182	VKAYVELHIEQGRVLEETGLPVGIVTGIAGLIWVKFTIEGKAEEHAGATPMSSLRRDEAAA	241						
Qy	246	ALMV---REVNRFVNEIADGTVATVGHILTAPGGGNQVPGEVDFTLDLRSPEEESLRV	301						

Qy 369 VTDVGMVFVPSRAGRSHVPEEWTDFDDLRKGTEVVLVRMK 408
 Db 366 LTEVGLIFVPSHNGISHAPEEWTDYDKLQKGIEVVLKTVK 405

RESULT 2

AE1500 N-carbamyl-L-amino acid amidohydrolase homolog lin0541 [imported] - Listeria innocua (strain

C;Species: Listeria innocua

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C;Accession: AE1500

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker ; Domínguez-Bernal, G.; Duchaude, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U.

A;Authors: Kreft, J.; Kuhn, M.; Kurapkat, G.; Madueno, E.; Maitourna, A.; Ma-

Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voß, H.; Wehland,

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AE1500

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-414 <GLA>

A;Cross-references: GB:AL592022; PIDN:CAC95773.1; PID:916412981; GSPDB:GN00178

A;Experimental source: strain Clip1.1262

C;Genetics:

A;Gene: lin0541

C;Superfamily: N-carbamyl-L-amino acid amidohydrolase

Query Match Score 41.6%; DB 2; Length 414;
 Best Local Similarity 43.2%; Pred. No. 7e-50;
 Matches 173; Conservative 73; Mismatches 152; Indels 2; Gaps 2;Qy 10 RIEKEIRELSRFSA-EQPGVTRLTYPHEAAARETLLIAAMKAALSVRDAGLNIIGRE 68
 Db 7 RIKKHIENLDRFTATPGQGTTTRLYSKEDLGARNYKEEMAKVGLTVSEDAIGNIYGRL 66Qy 69 GTDPPELPAIAVGSHFDSVRNGGMFDGTAGVVCALLEAARVMLENGYVNHRHPFEFIAIVEEE 128
 Db 67 GDNPDIPIAVVGSHFDSVPNGGAFDGPAGVITGLEVASVFEQQIKPHPFLEITIAMVEE 126Qy 129 GARFSSGMLGGRAIAGLVADREELDSLVDDEDGVSVRQAATAFGLKPGELQAAARSAAIDLRA 188
 Db 127 GSRFGAGLLASRTITGKVTKEMLDINGVTAEEAMANLGFDANQVHTAIRSKDSVKA 186Qy 189 FIELHIEQGPITLEEQQIEIGVVTTSIVGVRALRVAVKGRSDDHAGTTPMHLLRQDALVPAALM 248
 Db 187 FIELHIEQGPVLENANEDVALVDTVVGLETEIKVTVKGQAGHAGTTPMLERKDALSAAVQI 246

Qy 249 VREVNRFVNEIADGTATVGHLTVAPEGGNQVPGEVDFTLDLRSPHEEESLRVLIDRISVM 308

Db 247 LNKLPELAIQEGGGTVLTIKLNVYPNGANVNPKVFTVDIRAKDEIHVQNTLETRKVV 306

Qy 309 VGEVASQAGVAADVDEFNLSPVQALPTMVDAVREAASALQFTYRDISSGAGHDMSMFLAQ 368

Db 307 I-QAFERNGIMCEIEDMLYEKPTPLSTEIHQALTESADKGLKTYRTMVGAGHDAMIFAS 365

Qy 369 VTDVGMVFVPSRAGRSHVPEEWTDFDLRKGTTEVVLVRMK 408
 Db 366 LTEVGLIFVPSHNGISHAPEEWTDYDKLQKGIEVVLKTVK 405

RESULT 3

A83745 N-carbamyl-L-amino acid amidohydrolase BH0761 [imported] - Bacillus halodurans (strain

C;Species: Bacillus halodurans

C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001

C;Accession: A83745

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hirai

Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliophilic bacterium Bacillus halodurans and

A;Reference number: A83650; MUID:11058132

A;Accession: A83745

Query Match Score 35.5%; DB 1; Length 414;

Best Local Similarity 37.3%; Pred. No. 1.7e-41;
 Matches 152; Conservative 82; Mismatches 165; Indels 8; Gaps 5;Qy 1 MTLOKAQAAERIEKEIRELSRF-SAEGPGVTRLTYPHEAAARETLLIAAMKAALSVREDA 59
 Db 4 VTISK--ERLRLHIEQOLGEIGKTKDKGVQRSLSKDREATLLVSEWNRAGGLTVTHDH 60

Qy 60 LGNLTIGRREGTDPPELPAIAVGSHFDSVRNGGMFDGTAGVVCALLEAARVMLENGYVNRRHPE 119

Db 61 FGNLIGRKEGETPSLPSVMIGSHIDSVRNGKFDGVIGVLAGIEIVHAISEANVVHEHS1 120

Qy 120 EFAIAVEEEARFSSGMLGGRAIAGLVADRELDLSVDEGVSVRQA-ATAFGLKPGELQ 177

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-414 <STO>

A;Cross-references: GB:AP001509; PIDN:BA000004; NID:910173176; PIDN:BA04480.1; GSPDB:GN00

A;Experimental source: strain C-125

C;Genetics:

A;Gene: BH0761

C;Superfamily: N-carbamyl-L-amino acid amidohydrolase

Query Match Score 37.2%; DB 2; Length 414;

Best Local Similarity 39.6%; Pred. No. 8.2e-44;
 Matches 161; Conservative 82; Mismatches 153; Indels 11; Gaps 4;Qy 10 RIEKEIRELSRFSAEGP---GVTRLTYPHEAAARETLLIAAMKAALSVRDAGLNIIG 65
 Db 11 RLWRTLMELAEIGGTPSPIDVHGVTRLSLTKTELARQYVIDLMKDSGLSVDQVDAVGNII 70Qy 66 REGTDPPELPAIAVGSHFDSVRNGGMFDGTAGVVCALLEAARVMLENGYVNRRHPEFIAIV 125
 Db 71 KLEGTDLELPSPVMTGSHIDSVPHGGRFDGTGLVGLAIEAVRTMKEAGTKLKHHSIEIVSFT 130Qy 126 EEEGARFSGMLGGRAIAGLVADRELDSSLVDEDGVSVRQAATAFGLKPGELQAAARSAAAD 185
 Db 131 DEEGARFGAGFIGSKGMAGELTETTF-SLADDKGVTYREAFALAANLNPTLYKQAIRSDEQ 189Qy 186 LRAFIELHIEQGPITLEEQQIEIGVVTTSIVGVRALRVAVKGRSDDHAGTTPMHLRQDALVPA 245
 Db 190 IKAYIEMHIEQGKVLEEHDSLISIGVTDQPVWLDTLEGAADHAGATPMDMRKD---A 245Qy 246 ALMVERVNRFVNEIA--DGTVATVGHLTVAPEGGGNQVPGEVDFFTLDLRSRPHESLRLVLI 303
 Db 246 GLAMAEVLAATEISKEHQGVGTVGKOMSIEPGGVNIIIPGRACFSVDLRHTRKERRQHMVD 305Qy 304 RISVMVGEVASQAGQVAADVDEFNLSPVQLAPTMVDVAREASALQFTHRISSGAGHDS 363
 Db 306 DLHEQVEAICNQRGTVTYNDVKKEPATCSHEMVGLIDEVCTELNIRAMKMPGAGHDA 365Qy 364 MFIAQVTDGMVFVPSRAGRSHVPEEWTDFDDDLRKGTEVVLVRMKAL 410
 Db 366 LIMSKLAPIGMIFIRSKQGISHSPKEWSDAEDCKKGQTQVLLHTLMKL 412Qy 426 ALMVERVNRFVNEIA--DGTVATVGHLTVAPEGGGNQVPGEVDFFTLDLRSRPHESLRLVLI 303
 Db 426 GLAMAEVLAATEISKEHQGVGTVGKOMSIEPGGVNIIIPGRACFSVDLRHTRKERRQHMVD 305Qy 435 RISVMVGEVASQAGQVAADVDEFNLSPVQLAPTMVDVAREASALQFTHRISSGAGHDS 363
 Db 436 MFLIAQVTDGMVFVPSRAGRSHVPEEWTDFDDDLRKGTEVVLVRMKAL 410Qy 444 MFIAQVTDGMVFVPSRAGRSHVPEEWTDFDDDLRKGTEVVLVRMKAL 410
 Db 445 D42594 N-carbamyl-L-amino acid amidohydrolase (EC 3.5.1.-) hyuC [similarity] - Pseudomonas sp.
 C;Species: Pseudomonas sp.
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 01-Sep-2000
 C;Accession: D42594
 R;Watabe, K.; Ishikawa, T.; Mukohara, Y.; Nakamura, H.
 J;Bacteriol. 174, 962-969, 1992
 A;Title: Cloning and sequencing of the genes involved in the conversion of 5-substituted
 A;Reference number: A42594; PMID:1732229
 A;Accession: D42594
 A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-414 <WAT>

C;Superfamily: N-carbamyl-L-amino acid amidohydrolase

C;Keywords: hydrolase

Query Match Score 35.5%; DB 1; Length 414;

Best Local Similarity 37.3%; Pred. No. 1.7e-41;
 Matches 152; Conservative 82; Mismatches 165; Indels 8; Gaps 5;Qy 1 MTLOKAQAAERIEKEIRELSRF-SAEGPGVTRLTYPHEAAARETLLIAAMKAALSVREDA 59
 Db 4 VTISK--ERLRLHIEQOLGEIGKTKDKGVQRSLSKDREATLLVSEWNRAGGLTVTHDH 60

Qy 60 LGNLTIGRREGTDPPELPAIAVGSHFDSVRNGGMFDGTAGVVCALLEAARVMLENGYVNRRHPE 119

Db 61 FGNLIGRKEGETPSLPSVMIGSHIDSVRNGKFDGVIGVLAGIEIVHAISEANVVHEHS1 120

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: May 3, 2004, 18:18:20 ; Search time 20 Seconds
 (without alignments)
 1981.545 Million cell updates/sec

Title: US-10-045-063-2
 Perfect score: 2049
 Sequence: 1 MTLOKAQAMRIEKEIRELSR.....FDDLRKGTEVVRLVUMKALDR 412

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing First 45 summaries

Database : PIR_78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	858	41.9	423	AB1142	N-carbamyl-L-amino acid amidohydrolase
2	853	41.6	414	AE1500	N-carbamyl-L-amino acid amidohydrolase
3	762.5	37.2	414	AB3745	N-carbamyl-L-amino acid amidohydrolase
4	728	35.5	414	DA2594	N-carbamyl-L-amino acid amidohydrolase
5	726.5	35.5	409	JN0885	N-carbamyl-L-amino acid amidohydrolase
6	671	32.7	411	D64079	N-carbamoyl-L-beta-alanyl dipeptidase
7	642.5	31.4	430	AH0394	N-carbamoyl-L-beta-alanyl dipeptidase
8	639.5	31.2	427	C83591	N-carbamoyl-L-beta-alanyl dipeptidase
9	612.5	29.9	427	HB7571	N-carbamoyl-L-beta-alanyl dipeptidase
10	602.5	29.4	412	G70017	N-carbamoyl-L-beta-alanyl dipeptidase
11	562	27.4	409	AG2869	N-carbamoyl-L-beta-alanyl dipeptidase
12	562	27.4	415	A97646	N-carbamoyl-L-beta-alanyl dipeptidase
13	562	27.4	416	F75429	N-carbamoyl-L-beta-alanyl dipeptidase
14	554.5	27.1	414	H95999	N-carbamoyl-L-beta-alanyl dipeptidase
15	536.5	26.2	415	AE3457	N-carbamoyl-L-beta-alanyl dipeptidase
16	534.5	26.1	411	B90701	N-carbamoyl-L-beta-alanyl dipeptidase
17	534.5	26.1	411	E85551	N-carbamoyl-L-beta-alanyl dipeptidase
18	528.5	25.8	411	AI0567	N-carbamoyl-L-beta-alanyl dipeptidase
19	528.5	25.8	525	FB5227	N-carbamoyl-L-beta-alanyl dipeptidase
20	525.5	25.6	411	C64783	N-carbamoyl-L-beta-alanyl dipeptidase
21	525	25.6	401	F90446	N-carbamoyl-L-beta-alanyl dipeptidase
22	501	24.5	413	AH2933	N-carbamoyl-L-beta-alanyl dipeptidase
23	501	24.5	441	G98348	N-carbamoyl-L-beta-alanyl dipeptidase
24	466.5	22.8	338	T04891	N-carbamoyl-L-beta-alanyl dipeptidase
25	172.5	8.4	371	E69963	N-carbamoyl-L-beta-alanyl dipeptidase
26	171	8.3	400	B72650	N-carbamoyl-L-beta-alanyl dipeptidase
27	168	8.2	372	E83833	N-carbamoyl-L-beta-alanyl dipeptidase
28	148	7.2	412	F87624	N-carbamoyl-L-beta-alanyl dipeptidase
29	146.5	7.1	381	E75298	N-carbamoyl-L-beta-alanyl dipeptidase

ALIGNMENTS

RESULT 1
 AB1142 N-carbamyl-L-amino acid amidohydrolase homolog lmo0537 [Imported] - Listeria monocytogenes
 C;Species: Listeria monocytogenes
 C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001
 C;Accession: AB1142
 R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke, J.; Dominguez-Bernal, G.; Duchaude, E.; Durand, L.; Dussurgey, O.; Entian, K.D.; Fsihi, H.D.; Jones, L.M.; Karst, U.; Science 294, 849-852, 2001
 A;Authors: Kreft, J.; Kuhn, M.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Moekane, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voiss, H.; Wehland, A.;Title: Comparative genomics of Listeria species.
 A;Reference number: AB1077; PMID:21537279; PMID:11679669
 A;Accession: AB1142
 A;Status: preliminary
 A;Molecule type: DNA
 A;Cross-references: GB:NC_003210; PIDN:CAC98616.1; PID:916409913; GSPDB:GN00177
 A;Experimental source: strain EGD-e
 C;Genetics:
 A;Gene: lmo0537
 C;Superfamily: N-carbamyl-L-amino acid amidohydrolase

Query Match 41.9%; Score 858; DB 2; Length 423;
 Best Local Similarity 44.2%; Pred. No. 3.3e-50;
 Matches 177; Conservative 68; Mismatches 153; Indels 2; Gaps 2;

Qy 10 RIEKEIRELRSRFSA-EGPGVTRLTYPPEHAARETLIAAMKAALSVREDALGNIIGRRE 68
 Db 7 RIKKHIEKLDTFTATPGQQTTTRLYSKEDLDARNYLQEMAKVGLTVSEDAIGNIYGRLE 66

Qy 69 GTDPPELPATAVGSHFDTSVRNGGMFDGTAGVWCALEAARVMLENGVNRHPFFIAIVEE 128
 Db 67 GESPDLPAVIVGSHFDSPVNNGGAFDGPAGVITGLEAVASVFHEQQIKPHPLEIIAMVEE 126

Qy 129 GARESSGMUGGRAIAGLVADRELDSLVDDEGVSRQATAFGLKPGELQAARSAAIDLRA 188
 Db 127 GSRFGAGLLLASRAITGKVTTTEMHEMKDIDGITATDAMAKLGFNANVRHARITKESVKA 186

Qy 189 FIELHIEQGPILEQEIQVGVTTIVGVRALRVAVKGRSDDHAGTPMHLRDQDALVPAALM 248
 Db 187 FIELHIEQGPILENASEDVALVDTVGLTQIKVTVKQAGHAGTPMLDRKDALVSAVEI 246

Qy 249 VREVNRFVNEIADGTVATVGHLTVAPEGGNQVPGEVDFTLDLRSPEEESLRVLIDRISVM 308
 Db 247 LGQLPPELAQEGGGTVLTVGKUNVYPNGANVIPDKVFTDVDIRAKDIEHVQNTLAKTKEI 306

Qy 309 VGEVASQAGVAADVDEFNLSPVQLAPTMVDAVREAASALQFTHRDISSGAGHDMSMFIAQ 368
 Db 307 I-QATEKNGITCEIKDMIYQQPTHLSKEITHQALTESADQLGFKYRTMVSGAGHDAMIFAS 365

Wed May 5 09:02:05 2004

us-10-045-063-2.rapb

Page 8

Qy 291 RSPHEESLRVLIDRISVMGEVASQAGVAADVDEFNFNLSPVQLAPTMVDAVREAAASALQF 350
| : | : | : | : | : | : | : |
Db 1196 RALGAE-----VTVAVCDVADRAAVAA-----LLAGLPADA-----PLSAVPH 1233
| : | : | : | : | : | : |
Qy 351 THRDISGGAGHDSMFIAQ--VTDVGMVFVPSRAGRSHVPEEWTDFFDLRKGTTEV 402
| : | : | : | : | : | : | : |
Db 1234 T-----AGVAH-SMPIGETGLTVAEVFAGKVAGARHL-----DELTRGHDL 1274
| : | : | : | : | : | : |

Search completed: May 3, 2004, 18:27:53
Job time : 49 SECs

US-10-282-122A-50358
 Query Match Score 6.8%; Score 138.5; DB 12; Length 418;
 Best Local Similarity 28.1%; Pred. No. 0.00026;
 Matches 45; Conservative 17; Mismatches 67; Indels 31; Gaps 4;
 Qy 3 LQKAQAAIRIEKIRELSRFSAEPPGVTRLTYTPEHAAARETLIAAMKAALSVREDAL-- 60
 Db 138 ISQLOASNVSTIVSLS----GFTNRYYTTTHGVAASDWIAQQWKQLAGSRDVTVE 190
 Qy 61 -----GNIIGRREGTDPELPATAVGSHFDS-----VRNGGMFDGTAGVVCAL 103
 Db 191 QFTAGWPQKSVLITKGSDPAAGVVVIGGHLSTDSTVGRMSENTRAPGADDASGIALTE 250
 Qy 104 AARYMLENGVNRHPPEFIAIVEEEGARFSSGMLGGRAIA 143
 Db 251 ALRVILANRYQPRTLFIGYAAE----AGLLGSQAI 285

RESULT 14
 US-10-282-122A-62122
 ; Sequence 62122, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cherry
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-09-09
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO: 62122
 ; LENGTH: 876
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium avium
 US-10-282-122A-62122
 Query Match Score 6.1%; Score 124; DB 12; Length 876;
 Best Local Similarity 24.0%; Pred. No. 0.019;
 Matches 95; Conservative 42; Mismatches 167; Indels 92; Gaps 14;

Qy 26 PGVTRILTYTPEHAAARETLIAAMKAALSVREDALGN1IGRRE-GTDPELPATAVG--- 80
 Db 465 PGATTLDVQAKHAAQQELAAALTGAAARSADQRRELGTRDQLGATLAGLCG 524
 Qy 81 -SHFDSVRN-----GGMFDTAGVVCALAAARVMLENGYVNRPFFEF 122
 Db 525 DEPLDELLRSRLARLRAEHPEQPAAPTGGAD----LAARTELEALAQAACASE 576
 Qy 123 AIVEEEGARFSSGMLGGRAIAIGLVADRELDSLVD-EGVSVRQATAFGLKRGEL--- 176
 Db 577 CVTRRQAAATAAARLAETATAATVTLQNRDLTTRAELDAATGRLAAERASVGDEELAASAD 636
 Qy 177 -----QAAARSAADL-----RAFILELHIEQGPLQEQIEI 207
 Db 637 AALRAEQAERRVAELADALAAASPDSVAEELADATESESLRREVYEEAGALREVTEL 696
 Qy 208 GVVTSIVGVRALRVAVKGRSDHAGTTPMHILRQDALVPALMVREV-----NRFVNE 258
 Db 697 SVFGS-EGROGKLDAETEREHAASE-HSQVGRARRAAQLRSVMTRHRDTTRQRYV-E 752
 Qy 259 IADGTVATVGHLTVAPGGGNQVPGEVDFTLDLRSPHEESLRVLIDRISVMGEVASOAGV 318
 Db 753 PYRAELQRLGR---PVFGPTFEVDIDSCLCIRSRTLNGITVYESLS---GGAKEQLGI 805
 Qy 319 AADVDEFNLSPVQLAPTMVDAVREAASALQFTHRD 354
 Db 806 LARLAGALVAKERDAPVUUVD----DALGFTDPD 835

RESULT 15
 US-10-205-032-16
 ; Sequence 16, Application US/10205032
 ; Publication No. US20030113874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Farnet, Chris
 ; APPLICANT: Yang, Xianshu
 ; APPLICANT: Staffa, Alfredo
 ; APPLICANT: Staffa, Alfredo
 ; TITLE OF INVENTION: GENES AND PROTEINS FOR THE BIOSYNTHESIS OF ROSARAMICIN
 ; FILE REFERENCE: 3016-2US
 ; CURRENT APPLICATION NUMBER: US/10/205,032
 ; CURRENT FILING DATE: 2002-07-26
 ; NUMBER OF SEQ ID NOS: 39
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO: 16
 ; LENGTH: 1574
 ; TYPE: PRT
 ; ORGANISM: micromonospora carbonacea subspecies aurantiaca
 US-10-205-032-16

Query Match 6.0%; Score 123.5; DB 14; Length 1574;
 Best Local Similarity 24.2%; Pred. No. 0.049;
 Matches 100; Conservative 50; Mismatches 143; Indels 121; Gaps 21;

Qy 7 QAARIEKEIRELSRFSAEPPGVTRLTYTPEHAAARETLIAAMKAALSVREDALGN1 63
 Db 964 QPAELVEAVRD-ALTARGAEVRLVTPSTDRAACAALLTAAGGGATRVL----SL 1016
 Qy 64 IGRREGTDPELPATAVGSHFDSVRNGGMFDGTAGVVCALAAARV-MLENGYVNRPFFEF 122
 Db 1017 LGTDRRPHPDHPAVSGAAATLLL----TQAVADALPAARLWVVTRGAVSVGPGE-- 1067
 Qy 123 AIVEEEGARFSSGMLGGRAIAIGLVADRELDSLVD-EDGVSVRQATAFGLKPGELQAAARS 182
 Db 1068 -TADERQAO---VWGFGRVAELPLRTWGLVD----LPADA-- 1101
 Qy 183 AADLRAFILELHIEQGPLQEQIEIGVVTISIVGVRAVKGRSDHAGTTPMHLRQDAL 242
 Db 1102 -----DGPVWE-----AFDVLAGDEDQVALRGPGVGYG---RRRLRRAPA 1137
 Qy 243 VPAALMVRREVNRFVNIEADGTVATVGHLT-VAPCGG-----GNQVPGEVDFTLSDL 290
 Db 1138 LPAKRRYRP---RGTVLVTGGTGALGAHVARLAAAGGAHLVLTSGRADAPGAAGLVGEI 1195

Db 9 ELLPIGRSSASG-GYRRFAWTGADACRAWFKAQAEQRGLTYELDRNGNQWAHLG---- 62 ; NUMBER OF SEQ ID NOS: 78614
 Qy 71 DPEL-PAIAGSHFDVRNGGMFDGTAGVVCALEAARVMLENGVYNNRHPFEIAVEEG 129 ; SOFTWARE: PatentIn version 3.1
 Db 63 DPTAGDAVVTGSHLDSPDGGAFDGPLGVSSFAAILDELRARQTLAKPLALVNFGDEG 122 ; SEQ ID NO 47606
 Db 130 ARFSSGMLGGRAIAGLVADRELDSLVDGDGSVRQAATAFGLKPGELQAAARSAAADLRAF 189 ; LENGTH: 319
 Qy 123 ARFGLACVGSRLLAQLTVECAHRLTDAVGTLPOAMERAGYDAGTTGDPERLARIGAF 182 ; TYPE: PRT
 Db 190 IELHIEQPILEQQEIQGVVTSIVGVRAVKGRSDHAGTPMHLRQDALVPAALMV 249 ; ORGANISM: Burkholderia cepacia
 Qy 183 VELHVEQRALDLSGDRIGIASAIWPHGRWRFDFRGEANHAGTRTRLADRRDPMLSYAETV 242 ; US-10-282-122A-47606
 Qy 250 REVNRFPVNEIADGTVATVGHLTVAQPGGNQVPGEVDFTLRSRHEESLRVLIDRISWV 309 ;
 Db 243 LAARREAR--LAGAVATFGKIAVEPNVGNAIPSLVRGWLDSRAADQRTLDVVTGVEKAA 300 ;
 Qy 310 GEVASQAGVAADVDEFNLSPVQLAPTMVDAVREA-ASALQFTH-----RD--- 354 ;
 Db 301 REYADAHGIELDV-----VRESFTPVVEFDHALRDELARILGRDTGT 342 ;
 Qy 355 -----ISSGAGHDMSMFLIAQVTDVGMVFVPSRAGRSHVPEEWTFDDLRKGTEV 402 ;
 Db 343 GAGTGADLKVPVLGTTGAHDAGILSGTIPTAMLFVVRNPNTGSHSPAEEYAAEDDCVAGWTA 402 ;
 Qy 403 VLVRMKALDR 412 ;
 Db 403 LADVLEGGLAR 412 ;
 RESULT 13
 US-10-282-122A-50358 ; Sequence 50358, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; PRIORITY APPLICATION NUMBER: US20040029129A1
 ; CURRENT FILING DATE: 2003-02-20
 ; PRIOR APPLICATION NUMBER: US/10/282,122A
 ; PRIOR FILING DATE: 2003-02-20
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.
 ; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 ; FILE REFERENCE: ELITRA.034A
 ; CURRENT APPLICATION NUMBER: US/10/282,122A
 ; PRIORITY APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/230,335
 ; PRIOR FILING DATE: 2000-09-06
 ; PRIOR APPLICATION NUMBER: 60/230,347
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/267,636
 ; PRIOR FILING DATE: 2001-02-09
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; Remaining Prior Application data removed - See File Wrapper or PALM.
 ; NUMBER OF SEQ ID NOS: 78614
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 50358
 ; LENGTH: 418
 ; TYPE: PRT
 ; ORGANISM: Burkholderia mallei

; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 11009
 ; LENGTH: 380
 ; TYPE: PRT
 ; ORGANISM: Ferroplasma acidarmatus
 US-10-369-493-11009

Query Match 23.2%; Score 474.5; DB 15; Length 380;
 Best Local Similarity 31.4%; Pred. No. 2.8e-36;
 Matches 121; Conservative 62; Mismatches 187; Indels 15; Gaps 6;

Qy 23 AEGPGVTRLTYTPEHAAARETLIAAMKAALSVREDALGNITGR-REGTDPELPAIAGVS 81
 :|:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 7 SEENGVNRLALNEYDIQARKNLEDEMRAVKADIKQDDAGLIFGTLGSGND---NTAIGS 62

Qy 82 HFDHSVNRNGGMFDGTAGVCALEAARVM---LENGYVNVRHPFEFIAIVEEGARFSSGMLG 138
 :|:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 63 HMDSVPNNGGRFDGFYGVMSGMQLLKDLGSTLKNRKITA---IDFTNEEGARFQPSLLG 117

Qy 139 GRAIAGLVADRELDLSLVDEDGVSVRQAATAFGLKPGELQAAARSAAADLRAFIELHIEQGP 198
 :|:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 118 SGMTSTGVFTKEFAYSRKDSDGITEEALEKSGYM-GEEQNRVKN-QDITRYLELHIEQGP 175

Qy 199 ILEEQQEIEIGVVTISIVGVRALRVAVKGRSDHAGTTPMHRLQDALVPAALMREVNRFVNE 258
 :|:|||:|||:|||:|||:|||:|||:|||:|||:
 Db 176 VLELENYQVGIPKGIVTMVVVDNVSFTGESNQAGPTPMKVRRDAVLVAASRFTVAVRDMAKS 235

Qy 259 IADGTVATVGHLTVAPGGGNQVPGEVDFTLDRSPHEESLRVLIDRISVMVGEVASQAGV 318
 :|:|||:|||:|||:|||:|||:|||:|||:
 Db 236 SGKELTMTVGKINNFPNAYNTVIPGKVTMNLDRSPDKKTAEEYSKMAKEISEKIASEEGV 295

Qy 319 AADVDEFNLSPVQLAPTMVDAVREAASALQFTTHDISSGAGHDSMFFIAQVTDVGMVFVP 378
 :|:|||:|||:|||:|||:|||:|||:
 Db 296 PVDFKRNQWTETTIFEDMRKEIIKSCKSLNLKYKELYSWPGHDAQYMNRRVVPtamifip 355

Qy 379 SRAGRSHVPEEWTFDDLRKGTEVV 403
 :|:|||:|||:|||:|||:
 Db 356 SHNGRSHTREEYSSDQDLINGYSVL 380

RESULT 11
 US-10-156-761-11032
 ; Sequence 11032, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO
 ; APPLICANT: HATTORI, MASAHIRA
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; FILE REFERENCE: 249-262
 ; CURRENT APPLICATION NUMBER: US/10/156,761
 ; CURRENT FILING DATE: 2002-05-29
 ; PRIOR APPLICATION NUMBER: JP 2001-204089
 ; PRIOR FILING DATE: 2001-05-30
 ; PRIOR APPLICATION NUMBER: JP 2001-272697
 ; PRIOR FILING DATE: 2001-08-02
 ; NUMBER OF SEQ ID NOS: 15109
 ; SEQ ID NO 11032
 ; LENGTH: 413
 ; TYPE: PRT
 ; ORGANISM: Streptomyces avermitillis
 US-10-156-761-11032

RESULT 10
 US-10-369-493-11009

; Sequence 11009, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; PLANTS WITH IMPROVED PROPERTIES
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21

Qy 14 EIRELSRSAEAGPGVTRLTYTPEHAARETLLIAAMKAALSVREDALGN---IIGRREGT 70
 :|:|||:|||:|||:|||:|||:
 ; Query Match 20.0%; Score 409.5; DB 14; Length 413;
 ; Best Local Similarity 29.1%; Pred. No. 4.7e-30;
 ; Matches 125; Conservative 56; Mismatches 192; Indels 57; Gaps 9;

60 ELGKVSADASGYLERTFLSPAFMXAIDLIRKWMEDAGLRTQEEEXMGNVHGLVDGANGNAEA 119
 77 IAVGSHFDVRNGGMFDGTAGVVCALEEARVMLENGYVN--RHPFEFIAIVEEGARFSS 134
 120 VXLGSHMDTVGNAGMFDSLGLIVSAISAWKAMHVNNGKLQKLKRPEVIAFSDEEGVRFQT 179
 135 GMLGGRAIAGLVADELDSLVDEDGVSVRQAATAFGLKPGELQAAARSAAADLR---- 187
 180 TFLGSGAIAGILP GTTLE-ISDKREVMKDF----LKENSMDITEESLLKLKYDPKSIW 233
 188 AFIELHIEQGPILEQEIQEIGVVTTSIVGVRALRVAKGRSDHAGTTPMHLRQDALVPA-
 234 GYVEVHIEQGPVLEQVGFPPLGVVKGIAGQTRLKVTVRGSQGHAGTVPMSMRDPMAAAEE 293

247 -LMVRE----- Sequence 8 ; Sequence 229957, Application US/10424599
 294 QIVVLESLCKHPEEYLSDGHCSDDSTVKS LSSSLVCTVGEISTWPSASNVIPGQVTTVD 353

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 354 IRAIDDLGREAVIYDLSKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 413

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 414 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 461

RESULT 7 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223)B ; FILE NUMBER: US/10/424,599
 ; CURRENT APPLICATION NUMBER: US/10/424,599 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684 ; NUMBER OF SEQ ID NOS: 229957
 ; SEQ ID NO 229957 ; LENGTH: 448
 ; TYPE: PRT ; ORGANISM: Glycine max
 ; FEATURE: NAME/KEY: unsure
 ; LOCATION: (1) .(448)
 ; OTHER INFORMATION: unsure at all Xaa locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_49676C.1.pep
 US-10-425-114-43946 ; Query Match 24.6%; Score 504; DB 12; Length 448;
 ; Best Local Similarity 34.1%; Pred. No. 5.7e-39; Matches 135; Conservative 57; Mismatches 178; Indels 26; Gaps 8;
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53313)B ; FILE NUMBER: US/10/425,114
 ; CURRENT APPLICATION NUMBER: US/10/425,114 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128 ; NUMBER OF SEQ ID NOS: 43946
 ; SEQ ID NO 43946 ; LENGTH: 511
 ; TYPE: PRT ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700682144_FLI.pep
 US-10-425-114-43946 ; Query Match 25.5%; Score 522.5; DB 12; Length 511;
 ; Best Local Similarity 32.1%; Pred. No. 1.2e-40; Matches 131; Conservative 73; Mismatches 163; Indels 41; Gaps 8;
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53313)B ; FILE NUMBER: US/10/425,114
 ; CURRENT APPLICATION NUMBER: US/10/425,114 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128 ; NUMBER OF SEQ ID NOS: 43946
 ; SEQ ID NO 43946 ; LENGTH: 511
 ; TYPE: PRT ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 700682144_FLI.pep
 17 ELSRFSAEGPGVTRLTYTPHEAAARETLIAAMKAALSREDALGNIIGREGTDPELPA 76
 88 ELGKVSADASGYLERTFLSPASMKAILRKWMEDAGLRTWVDQMGNVHGRVDGANENAEA 147

77 IAVGSHFDVRNGGMFDGTAGVVCALEEARVMLENGYVN--RHPFEFIAIVEEGARFSS 134
 148 LLIGSHMDTVGDGMFDGLIGIVSAISAWKAMHVNNGKLQKLKRPEVIAFSDEEGVRFQT 207

135 GMLGGRAIAGLVADELDSLVDEDGVSVRQAATAFGLKPGELQAAARSAAADLR---- 187
 208 TFLGSGAIAGILP GTTLE-ISDKREVMKDF----LKENSMDITEESLLKLKYDPKSIW 261

188 AFIELHIEQGPILEQEIQEIGVVTTSIVGVRALRVAKGRSDHAGTTPMHLRQDALVPA-
 262 GYVEVHIEQGPVLEQVGFPPLGVVKGIAGQTRLKVTVRGSQGHAGTVPMSMRDPMAAAEE 321

247 -LMVRE----- Sequence 8 ; Sequence 229957, Application US/10424599
 290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

RESULT 8 ; Sequence 229957, Application US/10424599
 290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 389
 442 KKMEGDIQDEVPTLMSGAGHDAMAISHLTKVGMFLFVRCRGGISHSPQE 489

290 LRSPEEESLRVLIDRISVMVGEVASOAGVAADVDEFNLSPVQLAPTMVDAVREAA-SAL 348
 382 IRAIDDLGREAVIYDL SKQIYQICDKRSVSCIIEHKHDAVICIDSDLSSQLKSAAYSAL 441

349 QFTHRDIS-----SGAGHDSMFIAQVTDVGMVFVPSRAGRSHVPEE 3

; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 49502
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE: Glycine max
; OTHER INFORMATION: Clone ID: 701134260_FLI.pep
; US-10-425-114-49502

Query Match 30.6%; Score 628; DB 12; Length 472;
Best Local Similarity 36.4%; Pred. No. 1e-50;
Matches 149; Conservative 66; Mismatches 188; Indels 6; Gaps 5;

Qy 8 AARIEKEIRELRSRFS-AEGPGVTRLTYTPEHAAARETLIAAMKAALSVREDALGNIIGR 66
Db 61 AQALQNQIDEELAFAFSDSAPSVTRVLYTDKDVLGRRYVKQTQMELAGLSVREDALGNIIGR 108

Qy 67 REGTDPPELAIAVGSHFDSVRNGGMFDGTAGVYCALEAARVMLENGYVNVRHPFEFFIAVE 126
Db 121 WDGLEPELAAVATGSHIDAIPIYSKGKYDGVGVGLGAEIAIRVLKRSGFKPRRPLEVISFTS 168

Qy 127 EEGARFSSGMLGGRAIAGLV-ADRELDSSLVDEDGVSVRQAATAFGLKPKGE--LQAAARSA 183
Db 169 EEPTRGIGCLGSRLLAGSEDLANSLKTSTDIONISFLDAAGSAGYSKNEDDLSVFLKK 228

Qy 184 ADLRAFIELDHIEQGPILEQEQQIEIGVVTISIVGVRALRVAVKGRSDHAGTTPMHLRQDALV 243
Db 229 GTYSAFVELHIEQGPILEDEGISIGIVTAIAAPASLTVEFEGNGGHAGAALMPNRNDAGL 288

Qy 244 PAALMREVNRFVNEIAD-GTVATVGHLTIVAPGGGNQVPGEVDFTLDLRSPEEESLRVLI 302
Db 289 AASELLAVERHVLDGSIDTVGTGILELHPGAINSIPSKSCHKLEIDTRDIDEERRNKVV 348

Qy 303 DRISVMVGEVASQAGVAADVDEFNLSPVQLA-PTMVDVAREASALQFTHRDISSGAGH 361
Db 349 EKIHQSAIKITKTRGVKLSDFRVINQDPPAISDEAVIKAVETATEKELNLKSXLMSRRAYH 408

Qy 362 DSMFIAQVTDGMVFVPSRAGRSHVPEEWTFDDLRKGTEVVLRVMKAL 410
Db 409 DSLFMARLSPMGMIIFIPCYKGYSHKPEEFATIEDMSNGVVKVLALTAKL 457

RESULT 6
US-10-424-599-158434
; Sequence 158434, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 158434
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Glycine max
; NAME/KEY: unsure
; LOCATION: (1) .. (483)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_114084C.1.pep
; US-10-424-599-158434

Query Match 25.7%; Score 526.5; DB 12; Length 483;
Best Local Similarity 32.1%; Pred. No. 4.7e-41;
Matches 131; Conservative 74; Mismatches 162; Indels 41; Gaps 8;

Qy 17 ELSRFSAEGPVTRLTYTPEHAAARETLIAAMKAALSVREDALGNTIGRREGTDPELPPA 76
; LENGTH: 460
; ORGANISM: Glycine max
; FEATURE:

Qy 121 FIAIVEEGARFSSGMLGRAIAGLVADRELDLVLVEDDGVSVRQAATAFGLKPGELQAAA 180
 Db 121 FIAIVEEGARFSSGMLGRAIAGLVADRELDLVLVEDDGVSVRQAATAFGLKPGELQAAA 180

Qy 181 RSAADLRAFIELHIEQGPILQEQQIEIGVVTTSIVGTRALRVAKGRSDHAGTTPMHLRQD 240
 Db 181 RSAADLRAFIELHIEQGPILQEQQIEIGVVTTSIVGTRALRVAKGRSDHAGTTPMHLRQD 240

RESULT 3
 US-10-289-360-4
 ; Sequence 4, Application US/10289360
 ; Publication No. US20030109013A1
 ; GENERAL INFORMATION:
 ; APPLICANT: TAKENAKA, YASUHIRO
 ; APPLICANT: SUZUKI, SHUNICHI
 ; APPLICANT: ONISHI, NORIMASA
 ; APPLICANT: YOKOZEKI, KENZO
 ; TITLE OF INVENTION: DNA ENCODING HYDANTOINASE, DNA ENCODING N-CARBAMYL-L-AMINO ACID
 ; TITLE OF INVENTION: HYDROLASE, RECOMBINANT DNA, TRANSFORMED CELL, METHOD OF PRODUCIT
 ; TITLE OF INVENTION: METHOD OF PRODUCING OPTICALLY ACTIVE AMINO ACID
 ; FILE REFERENCE: 229195US0CONT
 ; CURRENT APPLICATION NUMBER: US/10/289, 360
 ; CURRENT FILING DATE: 2002-11-07
 ; PRIOR APPLICATION NUMBER: PCT/JP02/02173
 ; PRIOR FILING DATE: 2002-03-08
 ; PRIOR APPLICATION NUMBER: JP 2001-0650814
 ; PRIOR FILING DATE: 2001-03-08
 ; PRIOR APPLICATION NUMBER: JP 2001-298619
 ; PRIOR FILING DATE: 2001-09-27
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; TYPE: PRT
 ; ORGANISM: Microbacterium liquefaciens
 ; LENGTH: 412
 ; US-10-289-360-4

Query Match 83.4%; Score 1709; DB 14; Length 412;
 Best Local Similarity 81.1%; Pred. No. 1.9e-153;
 Matches 38; Mismatches 40; Indels 0; Gaps 0;

Qy 361 HDMSMFIAQVTDVGMVFVPSRAGRSHVPEWTFDDLRKGTEVVLRMKALDR 412
 Db 361 HDMSMFIAQITDVGMVFVPSRAGRSHVPEEWSFDDLRKGTDVVLHVVTALDR 412

RESULT 2
 US-09-950-772-6
 ; Sequence 6, Application US/09950772
 ; Patent No. US20020102713A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SUZUKI, Shunichi
 ; APPLICANT: ONISHI, No. US20020102713A1masa
 ; APPLICANT: YOKOZEKI, Kenzo
 ; TITLE OF INVENTION: 5-SUBSTITUTED HYDANTOIN RACEMASE, DNA CODING FOR THE RACEMASE, AN
 ; TITLE OF INVENTION: PROCESSES FOR PRODUCING OPTICALLY ACTIVE AMINO ACIDS
 ; FILE REFERENCE: 212318US0
 ; CURRENT FILING DATE: 2001-09-13
 ; PRIOR APPLICATION NUMBER: US/09/950, 772
 ; PRIOR FILING DATE: 2000-09-13
 ; PRIOR APPLICATION NUMBER: JP2000-278571
 ; PRIOR FILING DATE: 2001-03-08
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 412
 ; TYPE: PRT
 ; ORGANISM: Microbacterium liquefaciens
 ; US-09-950-772-6

Query Match 83.4%; Score 1709; DB 9; Length 412;
 Best Local Similarity 81.1%; Pred. No. 1.9e-153;
 Matches 38; Mismatches 40; Indels 0; Gaps 0;

Qy 1 MTLOKAQARIEKEIRELSRFSAEQGPGVTRLTYPTEHAARETLLAAMKAALSVRDEL 60
 Db 1 VTLQQARADRIEEELWTLSRFSVEGPGVTRLTYPTEHAAREVVAAMQRTGLSVHEDAL 60

Qy 61 GNIIGRREGTDPPELPATIAVGSHFDSVRNGGMFDGTAGVVCALEAARVMLENGYVNRRHPFE 120
 Db 61 GNIIGRREGSDPALPAIAFGSHFDSVRNGGMFDGTAGVVCALEAARVLQESGYVNRRHPFE 120

Query Match 83.4%; Score 1709; DB 14; Length 412;
 Best Local Similarity 81.1%; Pred. No. 1.9e-153;
 Matches 38; Mismatches 40; Indels 0; Gaps 0;

Qy 121 FIAIVEEGARFSSGMLGRAIAGLVADRELDLVLVEDDGVSVRQAATAFGLKPGELQAAA 180
 Db 121 VIAIVEEGTRFSSGMLGRAIAGLVSDADDTLVDEDGTVREATAFGLEPGEIRTA 180

Qy 181 RSAADLRAFIELHIEQGPILEQEQIEIGVVTTSIVGTRALRVAVKGRSVDHAGTTPMHLRQD 240
 Db 181 RTRDDLRAFIELHIEQGPILEQEQIEIGVVTGIVGVRAFRITVEGRSDHAGTTPMHLRQD 240

Qy 241 ALVPAALMVRENVRFVNEIADGTVATYGHLTVPAPGGGNQVPGEVDFTLDSLRSPEESLRV 300
 Db 241 ALVPAALMVRINRFVNEIADGTVATYGHLTVPAPGGGVFTDLRSPEESIRL 300

Qy 301 LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMVDAVREAASALQFTHRDITSGAG 360
 Db 301 LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMVDAVREAASALQFTHRDITSGAG 360

RESULT 4
 US-10-425-114-49502
 ; Sequence 49502, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: May 3, 2004, 18:22:06 ; Search time 48 Seconds
 (without alignments)
 2379.211 Million cell updates/sec

Title: US-10-045-063-2

Perfect score: 2049

Sequence: 1 MTLOQKAQAAERIEKEIRELSR.....EDDLRKGTTEVVLRVMKALDR 412

Scoring table: BLOSUM62

Gappen 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/us07_PUBCOMB.pep:*

2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/2/pubpaa/us06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/2/pubpaa/us06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/2/pubpaa/us07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*

7: /cgn2_6/ptodata/2/pubpaa/us08_NEW_PUB.pep:*

8: /cgn2_6/ptodata/2/pubpaa/us08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/2/pubpaa/us09A_PUBCOMB.pep:*

10: /cgn2_6/ptodata/2/pubpaa/us09B_PUBCOMB.pep:*

11: /cgn2_6/ptodata/2/pubpaa/us09C_PUBCOMB.pep:*

12: /cgn2_6/ptodata/2/pubpaa/us09_NEW_PUB.pep:*

13: /cgn2_6/ptodata/2/pubpaa/us10A_PUBCOMB.pep:*

14: /cgn2_6/ptodata/2/pubpaa/us10B_PUBCOMB.pep:*

15: /cgn2_6/ptodata/2/pubpaa/us10C_PUBCOMB.pep:*

16: /cgn2_6/ptodata/2/pubpaa/us10_NEW_PUB.pep:*

17: /cgn2_6/ptodata/2/pubpaa/us60_NEW_PUB.pep:*

18: /cgn2_6/ptodata/2/pubpaa/us60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2049	100.0	412	14	US-10-334-990-7
2	1709	83.4	412	9	US-09-950-772-6
3	1709	83.4	412	14	US-10-289-360-4
4	628	30.6	472	12	US-10-425-114-49502
5	625	30.5	460	12	US-10-424-599-229959
6	526.5	25.5	483	12	US-10-425-114-439434
7	522.5	25.5	511	12	US-10-424-599-229957
8	504	24.6	448	12	US-10-424-599-229957
9	476.5	23.3	380	15	US-10-369-493-10895
10	474.5	23.2	380	15	US-10-369-493-11009
11	409.5	20.0	413	14	US-10-156-761-11032
12	143.5	7.0	319	12	US-10-282-122A-47606
13	138.5	6.8	418	12	US-10-282-122A-50358
14	124	6.1	876	12	US-10-282-122A-62122
15	123.5	6.0	1574	14	US-10-205-032-16

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	2049	100.0%	412	14	US-10-334-990-7
2	1709	100.0%	9	6	US-09-950-772-6
3	1709	100.0%	14	14	US-10-289-360-4
4	628	100.0%	472	12	US-10-425-114-49502
5	625	100.0%	460	12	US-10-424-599-229959
6	526.5	100.0%	483	12	US-10-425-114-439434
7	522.5	100.0%	511	12	US-10-424-599-229957
8	504	100.0%	448	12	US-10-424-599-229957
9	476.5	100.0%	380	15	US-10-369-493-10895
10	474.5	100.0%	380	15	US-10-369-493-11009
11	409.5	100.0%	413	14	US-10-156-761-11032
12	143.5	100.0%	319	12	US-10-282-122A-47606
13	138.5	100.0%	418	12	US-10-282-122A-50358
14	124	100.0%	876	12	US-10-282-122A-62122
15	123.5	100.0%	1574	14	US-10-205-032-16

Result No.	Score	Query Match	Length	DB ID	Description
1	2049	100.0%	412	14	US-10-334-990-7
2	1709	100.0%	9	6	US-09-950-772-6
3	1709	100.0%	14	14	US-10-289-360-4
4	628	100.0%	472	12	US-10-425-114-49502
5	625	100.0%	460	12	US-10-424-599-229959
6	526.5	100.0%	483	12	US-10-425-114-439434
7	522.5	100.0%	511	12	US-10-424-599-229957
8	504	100.0%	448	12	US-10-424-599-229957
9	476.5	100.0%	380	15	US-10-369-493-10895
10	474.5	100.0%	380	15	US-10-369-493-11009
11	409.5	100.0%	413	14	US-10-156-761-11032
12	143.5	100.0%	319	12	US-10-282-122A-47606
13	138.5	100.0%	418	12	US-10-282-122A-50358
14	124	100.0%	876	12	US-10-282-122A-62122
15	123.5	100.0%	1574	14	US-10-205-032-16

Result No.	Score	Query Match	Length	DB ID	Description
1	2049	100.0%	412	14	US-10-334-990-7
2	1709	100.0%	9	6	US-09-950-772-6
3	1709	100.0%	14	14	US-10-289-360-4
4	628	100.0%	472	12	US-10-425-114-49502
5	625	100.0%	460	12	US-10-424-599-229959
6	526.5	100.0%	483	12	US-10-425-114-439434
7	522.5	100.0%	511	12	US-10-424-599-229957
8	504	100.0%	448	12	US-10-424-599-229957
9	476.5	100.0%	380	15	US-10-369-493-10895
10	474.5	100.0%	380	15	US-10-369-493-11009
11	409.5	100.0%	413	14	US-10-156-761-11032
12	143.5	100.0%	319	12	US-10-282-122A-47606
13	138.5	100.0%	418	12	US-10-282-122A-50358
14	124	100.0%	876	12	US-10-282-122A-62122
15	123.5	100.0%	1574	14	US-10-205-032-16

Result No.	Score	Query Match	Length	DB ID	Description
1	2049	100.0%	412	14	US-10-334-990-7
2	1709	100.0%	9	6	US-09-950-772-6
3	1709	100.0%	14	14	US-10-289-360-4
4	628	100.0%	472	12	US-10-425-114-49502
5	625	100.0%	460	12	US-10-424-599-229959
6	526.5	100.0%	483	12	US-10-425-114-439434
7	522.5	100.0%	511	12	US-10-424-599-229957
8	504	100.0%	448	12	US-10-424-599-229957
9	476.5	100.0%	380	15	US-10-369-493-10895
10	474.5	100.0%	380	15	US-10-369-493-11009
11	409.5	100.0%	413	14	US-10-156-761-11032
12	143.5	100.0%	319	12	US-10-282-122A-47606
13	138.5	100.0%	418	12	US-10-282-122A-50358
14	124	100.0%	876	12	US-10-282-122A-62122
15	123.5	100.0%	1574	14	US-10-205-032-16

Result No.	Score	Query Match	Length	DB ID	Description
1	2049	100.0%	412	14	US-10-334-990-7
2	1709	100.0%	9	6	US-09-950-772-6
3	1709	100.0%	14	14	US-10-289-360-4
4	628	100.0%	472	12	US-10-425-114-49502
5	625	100.0%	460	12	US-10-424-599-229959
6	526.5	100.0%	483	12	US-10-425-114-439434
7	522.5	100.0%	511	12	US-10-424-599-229957
8	504	100.0%	448	12	US-10-424-599-229957
9	476.5	10			

This Page Blank (uspto)

Best Local Similarity 22.5%; Pred. No. 0.047;
 Matches 106; Conservative 42; Mismatches 169; Indels 155; Gaps 21;

Qy 27 GVTRLTYPE----HAAARETLIAAMKA-----AALSVERDALGNIIGRREG--T 70
 Db 11127 GVDRFGVHPALLDAVLHAAAETSVVQSEARVPFSWRGVELRATESA---VYRARLSLTS 1183

Qy 71 DPELPATAVGSHFDSVRNGGMFDGTAGVVCALEARVMLENGVNRRHPPFEF----- 121
 Db 11184 DDELSLVAVDP-----AGRIVATVDSLVTTRPISRQQVRSGAIGDCLFEVEMHRKALLG 1236

Qy 122 -----IAIVEEEBEGARFSSGMLGGRAIAGLVADEFLDSLVDDED----GVSVRQAATAFGL 171
 Db 1237 TTAGDDLAIV-GDGPSWPESV---RATARFATLDEFRAAVDSDVPAPGSWVLAAMSAEV 1292

Qy 172 KPGELQAAAR-SPADLRAFILELHIEQGPITLEQEQUI-----GVVTSIYGV---- 216
 Db 1293 EGGSLPSRAQUESTSDLLLALVQSWLADERFAESOLVUVTRAAVSADSDDSDVADLVGASSWG 1352

Qy 217 -----RALRVAVKGRSDD-----HAGTTPMHLRQD-ALVP--AALMVREVN 253
 Db 1335 LLSSAQSENPGREVLDVDTGPESWQALPAAVRAAGEPQLALRGVALVPRLARLTVREG 1412

Qy 254 RFVNEIADGTVATVG-----HLTVAPGGGNQVPGEVDFTLDLRSP 293
 Db 14113 SSPQLDDGTVLITGGTGALGGVARHILVEEHGIRRLVLAGRRGNAPG-----V 1462

Qy 294 HEESLRVLID--RISVMVGEVASQAGVAADVDEFNLSP-----VQLAPTMVDAVR 342
 Db 1463 HE----LVDELALARAGAVVEVACDVADRTDLEHVLLAAIPVWDPLRGIVHTAGVLADGVI 1517

Qy 343 EAASALQFTHRDISSGAGHDSMFIAQVTDVGMOVFPVPSRAGRSHVPEEWTDFD 394
 Db 1518 GSLSA-----ADVGTVFAPKVTVGAWHLHELTRLD 1547

Search completed: May 3, 2004, 18:23:05
 Job time : 24 secs

FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5309
; LENGTH: 406
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
; US-09-328-352-5309

Query Match 6.5%; Score 133; DB 4; Length 406;
Best Local Similarity 21.0%; Pred. No. 7e-05;
Matches 88; Conservative 67; Mismatches 170; Indels 94; Gaps 19;

Qy 26 PGVTRLTYTPEHAARETLIA-AMKAAALSVRDAGNI--IGRREGTDPELPAIAVGSH 82
Db 46 PSVTPIDHT-----CQTIMADRLAKVGFHIEPMRFQGDVDNLWARRGT--EGPVFCFAHG 97
Qy 83 FD-----SVRNGGMFD-GTAGVVCALEAARVMLENGYVNHRP----F 119
Db 98 TDVVPTGRDLDAWNSDPFAPEI RDGKLYGRGSADMKTAL-AAMVVASERFVAKHPNHKGGSI 156
Qy 120 EFTIAIVEEGARFSSGMLGGRAIAGLVAADRELDLSVDEDGVSVRQAATAFGLKPGELOAA 179
Db 157 AFLITSDEEGP-----AVNGTVKVIETLEKRNEKITWCLVGEPSSTHKLDIVKN 206

RESULT 15
US-09-036-987A-5 ; Sequence 5, Application US/09036987A
; Patent No. 6143526 ; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Treadaway, Patti J.
; APPLICANT: Turner, Jan R.
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; STREET: Dow AgroSciences LLC Patent Department
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 337-4816
; TELEFAX: (317) 337-4847
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4928 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-036-987A-5

Query Match 6.2%; Score 127.5; DB 4; Length 1002;
Best Local Similarity 24.2%; Pred. No. 0.001;
Matches 101; Conservative 44; Mismatches 125; Indels 147; Gaps 23;

Qy 3 LQKAQARIEKEIRELSRFSAAE-----PGVTRLTYTPEHAAAR-----E 42
Db 460 LRAAQVARV-----LEGQPRVAGLEQHQGENLAPQVLGLDDLVQLQAVAGQGLV 508

Query Match 5.9%; Score 121.5; DB 3; Length 4928;

RESULT 11
 US-09-328-352-6356
 ; Sequence 6356, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: GTC99-03PA
 ; CURRENT APPLICATION NUMBER: US/09/328, 352
 ; CURRENT FILING DATE: 1999-06-04
 ; NUMBER OF SEQ ID NOS: 8252
 ; SEQ ID NO: 6356
 ; LENGTH: 447
 ; TYPE: PRT
 ; ORGANISM: Acinetobacter baumannii
 US-09-328-352-6356

Query Match 6.9%; Score 140.5; DB 4; Length 447;
 Best Local Similarity 20.8%; Pred. No. 1.4e-05;
 Matches 93; Conservative 74; Mismatches 174; Indels 107; Gaps 21;

Query 12 EKEIRELSRFSAAEGPGVTRLTYTPHEAAARETLIAMIKAALSVRED-ALGNIIGRREGT 70
 34 ENQVIQLORQHIEYYPELGNGMNEFKTSALWQKE----LKSYGIVQKTKGYAKTVGIGILGN 88

Query 71 DPELPAIAVGSHFDSV-----RNGGMFEDG-----TAGVVCALEAARV 107
 89 NPG-P11ALRADMDALPMEEKSGVPFASKQOKAIYQGKETYVMHACGHDAAHTMLLGAAKI 147

Query 108 MLENGYVNRRHPFEPFIAVEEGA---RFSSG-MLGGRAIAGLVADRELDLSVDED-GV 160
 Db 148 LAANKDKISGTFFFQPAEEGGADIDNFTHGDQIGSRK---MIADGAFKDYKPEAIFGM 204

Query 161 SVRQAAATAFGLKPGELQAAARSAAIDLRAFIELHIEQGPFILEQQIEGVVTSIVGVRALR 220

Db 205 HVMS----GMKSG-----HLYYKDGAIIINSAD-----HLR 230

RESULT 12
 US-09-252-991A-21048
 ; Sequence 21048, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: MARC J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196_136
 ; CURRENT APPLICATION NUMBER: US/09/252, 991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074, 788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094, 190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO: 21048
 ; LENGTH: 434
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-21048

Query Match 6.5%; Score 133.5; DB 4; Length 434;
 Best Local Similarity 20.5%; Pred. No. 6.9e-05;
 Matches 84; Conservative 53; Mismatches 154; Indels 119; Gaps 14;

Query 21 FSAEAGPGVTRLTYT----PEHAARETLIA-----AMKAALSVREDALG----61
 45 FAADTPAAELLRQAAEAERPAYLDTLRLQLVAVDSGTGQAEGLGLQSLALLAERIQLALGAQVR 104

Query 62 -----NIIGREGTDPPELPAIAVGSHFDSV-----RNNGMFD 93
 Db 105 SAPATPSAGDNLVATLDGTGSKRFLLMI--HYDTVFAAGSAAKRPFREDAERA YGPGVAD 162

Query 94 GTAGVVCALEAARVMLENGIVNRHPFEPFIAIVEEGARFSSGMGLGGRAIAGLVADRELDS 153

Query 163 AKGGVAMVHLALLRQQFRDGRITVLNFNPDEE----TGSAGSKQLIAELA-----211

Query 154 LVDEDGVSVRQAAATAFGLKPGELQAAARSAAIDLRAFIELHIEQGPFILEQQIEGVVTSI 213

Query 212 -----RQDYYFSYEP PDRDA-----VTVAT 232

Query 214 VGVRALRVAVKGRSRSDHAGTTPMHLRQDALVPAALMVREVNRFVNNEIADGTVATVGHLTVA 273

Query 233 NGIDGLLLEVKGRRSHAGSAPEGRNAILESHOLLR-----LKDLGDPAKGTTLNWTLA 287

Query 274 PGG--GNQYPGEVDFTLDLR-SPEEESLRVLIDRISVMGEVASQAGVAADVDEFNLSP 330

Query 288 RGGEKRNIIPAEASAADMRYSDPASERVLA DARKLTGERLVADETEVSLRLDKG----342

Query 331 VQLAPTMVDAVREAAASALQFTHDISSGAGHDSMFIAQ---VTDVGMVFVP 378

Query 343 ---RPPLVK--NPASQRLAETAQFLYGRVGRIEPIAMRFGTDAGYAYVP 387

RESULT 13
 US-09-328-352-5309
 ; Sequence 5309, Application US/09328352
 ; Patent No. 6562958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary L. Breton et al.
 ; TITLE OF INVENTION: NUCLEAR ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
 ; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
 ; TITLE OF INVENTION: TAGNVIPEQVNMIIGTIRSNNEQIRENILKSLPALVEHNAQANEVTAKV-EIAPYAPVTMN 344

PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12025
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-12025

Query Match 20.9%; Score 428; DB 4; Length 307;
Best Local Similarity 35.8%; Pred. No. 1.1e-34;
Matches 102; Conservative 51; Mismatches 120; Indels 12; Gaps 3;

Query Match 7.0%; Score 143.5; DB 4; Length 465;
Best Local Similarity 20.3%; Pred. No. 7.6e-06;
Matches 94; Conservative 73; Mismatches 173; Indels 123; Gaps 19;

Qy 127 EEEGARFSSGMLGGRAIAIGLVAADRELDLSVDEDGVSYRQAATAFGLK----PGELQAAARS 182
Db 23 EEGVRFPQDMMGSVVFFAGEYPLAQALAAKDLGDTIDEALRNIGYKGGERQPGDM---- 76

Qy 183 AADLRAFILEHIEQPILEQEQQIEGVVTSIVGVRALRVAKGRSDDHAGTPMHLRQDAL 242
Db 77 AVD--SYVELHIEQGPILDKEQIDIGVVTGVQQGISWQEFTLRGVSNHAGTPMSMRDAG 134

Qy 243 VPAALMVREVNRFVNIEADGTVATVGHLTVAAPGGGNQVPGEVDFTDLRSPEESLRVLI 302
Db 135 LAAAIAKIAVFARELLAISLGGNQVATVGHFSVKPNLINVIPNHHVMSVDLRNTDNAILCLAE 194

Qy 303 DRISVMVGEVASQAGVAAADVDEFFNLSPVQLAPTMVDAVREASALQFTTHDISSGAGHD 362
Db 195 QQLAEFFVAKTQSQEGVEITSRSIYRFNPVIADEIVNAVEAEERQALSYRRLLPSGAGHD 254

Qy 363 SMFIAQVTDVGMFVFVPSRAGRSHVPEEWTFDDLRLKGTEVVLRM 407
Db 255 AQFMASVCPAGMIFVPCVDFGISHNVKEHSAAKDLIAGANVLLQQV 299

RESULT 8
US-09-489-039A-12144
; Sequence 12144, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12144
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-12144

Query Match 8.0%; Score 163.5; DB 4; Length 134;
Best Local Similarity 39.4%; Pred. No. 1.1e-08;
Matches 37; Conservative 15; Mismatches 39; Indels 3; Gaps 2;

Qy 25 GPGVTRLTYPHEAAARETLIAAMKAALSVREDALGNIIGRREGTDPELPAIAVGSHFD 84
Db 35 GGGVCRLLAATAEDKAGRDFVARMKALGLNVNSIDAIGNVTGVYHG-EETLPMVMMGSHID 93

Qy 85 SVRNGGMEDGTAGVVCALEAARVMLENGVNRHP 118
Db 94 TFATDGLYDGKYGMAGLEVIAITFQDAG--DPHP 125

RESULT 9
US-09-328-352-7315
; Sequence 7315, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et. al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

Query Match 6.9%; Score 141.5; DB 4; Length 383;
Best Local Similarity 19.5%; Pred. No. 9e-06;
Matches 77; Conservative 67; Mismatches 138; Indels 113; Gaps 15;

Qy 58 DALGN--TIGRREGTDPELPALIAVGSHFDSV-----RNGGMFDGTAG 97
; ORGANISM: Proteus mirabilis
; US-09-543-681A-7526

RESULT 10
US-09-543-681A-7526
; Sequence 7526, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7526
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; US-09-543-681A-7526

Qy 60 LGNIIIGRREGTDFELPAIAVGSHFDFSVERNGGMFDGTAGVYCALEAARVMILENGYVNRIHPF 119
 Db 75 VGNLFCRVAGTERPQETILTGSHIDTVNGGTLGQGYVIASMTAVQYLLDKYKPRSL 134
 Qy 120 EFAIAVEEGARFSSGMLGGRAIAGL-----VADRELDSSLYDE---DGVSVRQATA 168
 Db 135 EVLSMAEEEGSRPPTVFWGSKNVGEAKREEVVDISDNEGKFKVDEMHRQGPDFRQEQT- 193

RESULT 6
 US-09-543-681A-5198
 ; Sequence 5198, Application US/09543681A
 ; Patent No. 6605709
 ; GENERAL INFORMATION:
 ; APPLICANT: GARY BRETON
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
 ; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.1002-001
 ; CURRENT APPLICATION NUMBER: US/09/543,681A
 ; CURRENT FILING DATE: 2000-04-05
 ; PRIORITY APPLICATION NUMBER: US 60/128,706
 ; PRIOR FILING DATE: 1999-04-09
 ; NUMBER OF SEQ ID NOS: 8344
 ; SEQ ID NO 5198
 ; LENGTH: 429
 ; TYPE: PRT
 ; ORGANISM: Proteus mirabilis
 ; US-09-543-681A-5198

Query Match 28.2%; Score 578.5; DB 4; Length 429;
 Best Local Similarity 33.2%; Pred. No. 1.4e-49;
 Matches 128; Conservative 66; Mismatches 188; Indels 3; Gaps 2;

Qy 229 HAGTTPMHLQRDALVPAALMREVNRFVNEIADGTVATVGHLTVAPEGGNQVPGEVDFTL 288
 Db 241 HAGTTPMHSYRHDAVYGFAKICAAEAIDRANEIGDPLVLTFGKVIPKPNTVNVPGEVEFTI 300
 Qy 289 DLRSPEEESLRVLIDRISVMGEVASQAGVAADVDEFNLSPVQLAPTMVDVREAAASL 348
 Db 301 DCRHTTDAFLRHFTEGEERMKTTAQELGLTIAIDRWMDAAPPVMNAIVEVIEQKAKEG 360
 Qy 349 QFTHREDISSGAGHDMSMFIAQVTDVGMVFVPSRAGRSHVPEWTFDDLRKGTEVV 403
 Db 361 KYQYRMHHSGAGHDSQITAPPHTAMIFVPSIGGISHNPAAEATAFPDLVEGVKL 415

RESULT 5
 US-09-554-998A-2
 ; Sequence 2, Application US/09554998A
 ; Patent No. 6599731
 ; GENERAL INFORMATION:
 ; APPLICANT: Estell, David A.
 ; TITLE OF INVENTION: Proteases from Gram Positive Organisms
 ; FILE REFERENCE: GC522-US
 ; CURRENT APPLICATION NUMBER: US/09/554,998A
 ; CURRENT FILING DATE: 2000-05-23
 ; PRIOR APPLICATION NUMBER: PCT/US98/27020
 ; PRIOR FILING DATE: 1998-12-17
 ; PRIOR APPLICATION NUMBER: GB 9727464.1
 ; PRIOR FILING DATE: 1997-12-30
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: FastSEQ for Windows Version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 412
 ; TYPE: PRT
 ; ORGANISM: Bacillus subtilisin
 ; US-09-554-998A-2

Query Match 29.4%; Score 602.5; DB 4; Length 412;
 Best Local Similarity 35.2%; Pred. No. 5e-52;
 Matches 137; Conservative 70; Mismatches 179; Indels 3; Gaps 2;

Qy 22 SAEGPCVTRLTYPEHAAARETLLIAAMKAAALSVREDALGNIIGRREGTDELPALAVGS 81
 Db 25 SADG-GVTRLTYKEWMDAQLAVKTEMSSGLETRFDVGNVFRGLSGTQSPDEVITGTS 83
 Qy 82 HFDSVRNGGMFDGTAGVYCALEAARVMILENGYVNRIHPFEFIAIVEEGARFSSGMLGGRA 141
 Db 84 HIDTVINGKYDAGVLAAMLLALKQKLKETYGAPKTKLEAVSILCEESEGSRFPMTYWGSGN 143

Qy 142 IAGLVADRELDLSQLVEDDGVSVRQAAATAGLKPGELOQAARSAAIDLRAFIELHIEQGPL 201
 Db 144 MTGVFSSEQDAKEPRDESGVSLQTMHESGFGKGVFOQAYRT--DISAFVELHIEQGKTLE 201
 Qy 202 QEQIEIGVVTISIVGVRALRVAKGRSDHAGTTPMHLRQDALVPAALMIREVNRFVNEIAD 261
 Db 202 MSGRDIGIVTISIAGQRRLYLTLEGECNHAQTTSMKWKRKDPLAASSRIIHELLRSDELPD 261
 Qy 262 GTVATVGHLTVAPEGGNQVPGEVDFTLDRSPHEESLRVLIDRISVMGEVASQAGVAAD 321
 Db 262 ELRLTCGKITAEPNVANVIPGRVQFSIDIRHQHQHVALINGCLOKGIRAV 321
 Qy 322 VDEFFNLSPVQLAPTMVDVREAAASALQFTHRDISSGAGHDMSMFLAQVTDDGMVFVPSRA 381

RESULT 7
 US-09-489-039A-12025
 ; Sequence 12025, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIORITY APPLICATION NUMBER: US 60/117,747

RESULT 2
 US-09-489-039A-9613
 ; Sequence 9613, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/09/489, 039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117, 747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 9613
 ; LENGTH: 459
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-9613
 ; Sequence 9613, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/09/489, 039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117, 747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 9613
 ; LENGTH: 459
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae

Query Match 31.2%; Score 639.5; DB 4; Length 495;
 Best Local Similarity 38.1%; Pred. No. 1.3e-55;
 Matches 54; Mismatches 190; Indels 9; Gaps 3;

Query 2 TLQKAQAARIEKEIRELSRFSAEQP-GVTRLTYPPEHAAARETLIAAMKAALSVERDAL 60
 Db 79 TQRHIDGQRWLQSLMDLARLGATKGGYCRALSDLRQARDLFVQWCEAAGCTSVDRV 138

Query 61 GNIIGRREGTDPELPALPAIAVGSHFDSVRNGGMFDGTAGVVCALEARVMLENGVNRHPFE 120
 Db 139 GNIFARRPGRNPDLPPLPPVMTGSHIDTQPTGGKFDFGCFGWMAGLEVIRTLNDLGVETEAPLE 198

Query 121 FIAIVEEGARFSSGMLGGRAIAGLVADRELDSSLVDEDGVSVROATAFGLKPGELQAAA 180
 Db 199 VVWTNEEGSRFAPCNMGSGVFAFKFTLEETLAKRDADGVSVGEALDAIG----YAGA 252

Query 181 RSAA--DLRRAFTELHIEQGPILQEQQIEIGVVTISIVGWRALRAVKGRSDHAGTTTPMHLR 238
 Db 253 RDCLGHGPVGAYPEAHIEQGPILEDEEKTIGVVLGALQKWFDSLRLGVEAHAGGPTTPMHLR 312

Query 239 QDALVPAALMVRVNRFVNENIADGTVATVGHLTVAPEGCGGNQVPGEVDFTLDRSPHEESL 298
 Db 313 KDALVGAIAAVVEAVNRAALGHOPHACGTGTVGCLHAYPGSRNVIPEGVKMTLDFRHLQPERL 372

Query 4 QKAQAARIEKEIRELSRFSAEQP-GVTRLTYPPEHAAARETLIAAMKAALSVERDALGNI 63
 Db 51 REAAASRVMARADRLLAASFSETADALTRVYLSPEHQLQANQLVQGQWMDQDSVGN 110

Query 64 IGRREGTDPELPALIAVGSHFDSVRNGGMFDGTAGVVCALEARVMLENGVNRHPFEITA 123
 Db 111 CGRYEGQQEGAPAVLGLGSHLDTVRNAGRYDGMGLGVIAAEIVVQRLHQHGRRLAKAIEIVG 170

Query 124 IVEEEGARFSSGMLGGRAIAGLVADRELDSSLVDEDGVSVROATAFGLKPGELQAAAARS 183
 Db 433 AGHDAIFIFLAELGPAGMIFVPCENGISHNEIENASPDDLLAACAVLRLRAM 481

Db 171 FGDEEGTRFGITLGLSRGVGTGWPESWL-SQCDTIDGVSVQAQLVNAGLDPARIAHARHP 229

RESULT 4
 US-09-134-000C-5976
 ; Sequence 5976, Application US/09134000C
 ; Patent No. 6617156

Query 184 ADLRAFIETLHIEQGPILQEQQIEIGVVTISIVGWRALRAVKGRSDHAGTTPMHLRQDALV 243
 Db 230 RDIAYYLELHIEQGPICLEQAGLALGVVEAINGARRLNCRFTGEAGHAGTVPMLHRKDALA 289

Query 244 PAALMVREVNRFVNENIADGTVATVGHLTVAPEGCGGNQVPGEVDFTLDRSPHEESLRLVID 303
 Db 290 AAAEWMVQVENLTRQRGGNLVATVGTLCAPGAVNTVPEVQLTDIRGPQDAPLTALLE 349

Query 304 RISVMGEVVAQGVAADVDEFNNLSPVQLAPTMVDAVREAASALQFTHDISSGAGHDS 363
 Db 350 ELLGQQAQIAGRRLSFAAEYYRIAATAACDSHLQAVLSEAVVAVQGRSLTTPSGAGHDA 409

Query 364 MFIAQVTDVGMVFVPSRAGRSHVPEE 389
 Db 410 IAIAERWPSAMLFVRCLGGVSHHHPAE 435

RESULT 3
 US-09-252-991A-23393
 ; Sequence 23393, Application US/09252991A
 ; Patent No. 6551795

GENERAL INFORMATION:
 ; APPLICANT: MARC J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252, 991A

Query Match 29.6%; Score 606; DB 4; Length 425;
 Best Local Similarity 35.4%; Pred. No. 2.3e-52;
 Matches 68; Mismatches 170; Indels 30; Gaps 5;

Query 1 MTLOKAQAARIEKEIRELSRFSAE-GPVTRLTYPPEHAAARETLIAAMKAALSVERDA 59
 Db 19 MDLEKVLKQRIE---ELSAIGSDPAGGMTRLLYTDWLAQKYVQSOTEAFGLETOQFDE 74

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 18:19:20 ; Search time 22 Seconds
 (without alignments)
 966.814 Million cell updates/sec

Title: US-10-045-063-2
 Perfect score: 2049
 Sequence: 1 MTLQKAQAAARIEKEIRELSR.....FDDLRKGTTEVVLRVMKALDR 412
 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5
 Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Issued Patents AA:
 1: /cggn2_6/ptodata/2/iaa/5A_COMB.pep:
 2: /cggn2_6/ptodata/2/iaa/5B_COMB.pep:
 3: /cggn2_6/ptodata/2/iaa/6A_COMB.pep:
 4: /cggn2_6/ptodata/2/iaa/6B_COMB.pep:
 5: /cggn2_6/ptodata/2/iaa/PCTUS_COMB.pep:
 6: /cggn2_6/ptodata/2/iaa/backf1es1.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	2049	100.0	412	4	US-09-285-055-2	Sequence 2, Appli
2	640.5	31.3	459	4	US-09-489-039A-9613	Sequence 9613, Ap
3	639.5	31.2	495	4	US-09-252-991A-23393	Sequence 23393, A
4	606	29.6	425	4	US-09-134-000C-5976	Sequence 5976, Ap
5	602.5	29.4	412	4	US-09-554-998A-2	Sequence 2, Appli
6	578.5	28.2	429	4	US-09-543-681A-5198	Sequence 5198, Ap
7	428	20.9	307	4	US-09-489-039A-12025	Sequence 12025, A
8	163.5	8.0	134	4	US-09-489-039A-12144	Sequence 12144, A
9	143.5	7.0	465	4	US-09-328-352-7315	Sequence 7315, Ap
10	141.5	6.9	383	4	US-09-543-681A-7526	Sequence 7526, Ap
11	140.5	6.9	447	4	US-09-328-352-6356	Sequence 6356, Ap
12	133.5	6.5	434	4	US-09-252-991A-21048	Sequence 21048, A
13	133	6.5	406	4	US-09-328-352-5309	Sequence 5309, Ap
14	127.5	6.2	1002	4	US-09-252-991A-19662	Sequence 19662, A
15	121.5	5.9	4928	3	US-09-036-987A-5	Sequence 5, Appli
16	121.5	5.9	4928	3	US-09-370-700-5	Sequence 5, Appli
17	121.5	5.9	4928	4	US-09-603-207-5	Sequence 5, Appli
18	120	5.9	752	4	US-09-252-991A-21724	Sequence 21724, A
19	119	5.8	496	4	US-09-252-991A-17357	Sequence 17357, A
20	119	5.8	738	4	US-09-252-991A-30721	Sequence 30721, A
21	117.5	5.7	573	4	US-09-252-991A-31661	Sequence 31661, A
22	117	5.7	879	4	US-09-252-991A-28118	Sequence 28118, A
23	116.5	5.7	587	4	US-09-252-991A-31749	Sequence 31749, A
24	115	5.6	573	4	US-09-252-991A-20876	Sequence 20876, A
25	115	5.6	1151	4	US-09-252-991A-21328	Sequence 21328, A
26	114.5	5.6	471	4	US-09-252-991A-27633	Sequence 27633, A
27	113.5	5.5	820	4	US-09-252-991A-32001	Sequence 32001, A

ALIGNMENTS

RESULT 1	US-09-285-055-2	;	Sequence 2, Application US/09285055
		;	Patent No. 6352848
		;	GENERAL INFORMATION:
		;	APPLICANT: ALTENBUCHNER, JOSEF
		;	APPLICANT: MATTES, RALF
		;	APPLICANT: PIETZSCH, MARKUS
		;	APPLICANT: SYLDATK, CHRISTOPH
		;	APPLICANT: WIESE, ANJA
		;	APPLICANT: WILMS, BURKARD
		;	TITLE OF INVENTION: RECOMBINANT L-N-CARBAMOYLASE FROM ARTHROBACTER
		;	TITLE OF INVENTION: AURESCINS AND METHOD OF PRODUCING L-AMINO ACIDS
		;	FILE REFERENCE: RECOMBINANT L-N-CARBAMOYLASE
		;	CURRENT APPLICATION NUMBER: US/09/285,055
		;	CURRENT FILING DATE: 1999-04-02
		;	EARLIER APPLICATION NUMBER: DE 198 14 813.5
		;	EARLIER FILING DATE: 1998-04-02
		;	NUMBER OF SEQ ID NOS: 2
		;	SOFTWARE: PatentIn ver. 2.0
		;	SEQ ID NO 2
		;	LENGTH: 412
		;	TYPE: PRT
		;	ORGANISM: Arthrobacter aurescens
	US-09-285-055-2	;	
		;	Query Match
		;	Best Local Similarity 100.0%; Pred. No. 2.5e-197; Mismatches 0; Conservative 0; Indels 0; Gaps 0;
		;	Matches 412;
		;	MTLOKAQAAARIEKEIRELSRMLGGRGSHFDNSVRQAAARVMLENGIVNRHPFE 60
		;	MTLOKAQAAARIEKEIRELSRMLGGRGSHFDNSVRQAAARVMLENGIVNRHPFE 120
		;	GNIIGRREGTDPELPAPIAVGSHFDNSVRQAAARVMLENGIVNRHPFE 120
		;	MTLOKAQAAARIEKEIRELSRMLGGRGSHFDNSVRQAAARVMLENGIVNRHPFE 60
		;	GNIIGRREGTDPELPAPIAVGSHFDNSVRQAAARVMLENGIVNRHPFE 60
		;	FIAVEEEGARFSSGMLGGRAIAGLVADRELDLSVDEDGVSVRQAAAFGLKPGELQAAA 180
		;	FIAVEEEGARFSSGMLGGRAIAGLVADRELDLSVDEDGVSVRQAAAFGLKPGELQAAA 180
		;	RSAADLRAPIELHIEQPILEQQIEIGVYTSIVGVRAVKGRSDHAGTTPMHLRQD 240
		;	RSAADLRAPIELHIEQPILEQQIEIGVYTSIVGVRAVKGRSDHAGTTPMHLRQD 240

PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX DR WPI; 2001-639362/73.
 DR N-PSDB; AAS89217.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PR biodiversity.
 XX PS Claim 20; SEQ ID NO 55389; 103PP; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase Chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 502 AA;

Query	Match	Score	Length
Qy	48 MKAAALSVREDALGNITIGRREGTDPELPAIVGSHFDSVRNGGMFDGTAGVVCALEARRV	146;	502;
Db	1 MAASGLETRFDEVGNLYGRLNGTEYPOEVVLSGSHIDTVNGGNLDQFGALAAWLAIDW	9.2e-05;	107
Qy	108 MLENGYVNRPFPFFIAIVEBEGARFSSGMLGGRAIAGLVADRELDLVD	0;	156
Db	61 LKTQYGAPLRTVEVVTMAEEEGSRFPYFWGSKNTFGLANPDDVRNICD	0;	109

Query Match Score 146; DB 4; Length 502;
 Best Local Similarity 33.0%; Bred. No. 9.2e-05;
 Matches 36; Conservative 16; Mismatches 57; Indels 0; Gaps 0;

Search completed: May 3, 2004, 18:20:31
 Job time : 62 secs

CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations in
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC [ftp.wipo.int/pub/published_pct_sequences](ftp://wipo.int/pub/published_pct_sequences)

XX SQ Sequence 554 AA;

Query Match 17.1%; Score 349.5; DB 4; Length 554;
 Best Local Similarity 25.1%; Pred. No. 5.7e-24;
 Matches 115; Conservative 70; Mismatches 185; Indels 89; Gaps 13;

QY 9 ARIEKEIRE-----LSRFSAEAGPGVTRLT----YTPHEAAA---RETLIAA 47
 Db 111 ARIPQEVEKKEYTVLKSAPVYLEILDKRVNKGTGVKLADVLGIKPEEIMAIGDQENDIAM 170

QY 48 MKAAALSVREDALGNITIGRREGTDPELPATAVGSHFDSVRNGGMFDGTA----- 96
 Db 171 IEYAGVG-----AMDNAIPSVKEVANF--VTKSNLNEGEDGVAFAIEKPKDRRSA 215

QY 97 -GUVCALEAARV--MLENGYVNRHPPFEFIAIVEEGARFSSGMLGGRAIAGLVADREL D 152
 Db 216 SGTLAWLAPLNIDIRRFOGVLFVNKRKTAREFNGVAHQREQ---LVGSNRIFNV---KVN 268

QY 153 SLVDE--DGVSV-----ROQATAFGLKGEL-----QAAARSA 183
 Db 269 SVITQKVSSSGVTLYADTTKTFGMNRTILYVIDISSELTQRVISHVEEEAKIDDAEVHFL 328

QY 184 ADLRAFIELHIEQGPITLEQEQIEIGVVTISIVGVRALRVAVKGRSDDHAGTTPMHLRQDALV 243
 Db 329 TDIKAFVELHIEQGCVLESNGQSIGVVNAIVGQRRRTVTLNGESNHAGTTPMGYRRDTVY 388

QY 244 PAALMVREVNRFVNEIADGTVATVGHLLVAPGGGNQVPGEVDFTLDLRSPEEESLRVLIID 303
 Db 389 AFSRICHOSVEKAKRMDPVLTFGKVEPRPRNTVNVPGKTTFTIDCRHTDAAVLRFDTQ 448

QY 304 RISVMYGEVASQAGVAADVDEFNLSPVQLAPTMVDAVREAASALQFTHRDIISSGAGHDS 363
 Db 449 QLENDMRAICDEMIDIGIDIDLMDEEPVPMNKELVATLTELCECEREKLNYRVMHSGAGHDA 508

QY 364 MFIAQVTDVGMYFVPSRAGRSHVPEEWTDFFDLRKGTTEV 402
 Db 509 QIFAPRVRPTCMIFIPSINGISHNPAERTNITDL--GTEI 545

RESULT 14
 ABG08894 ID ABG08894 standard; protein; 435 AA.
 XX AC ABG08894;
 XX DT 13-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #8885.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US008631.
 XX PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX DR N-PSDB; AAS73081.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX PS Claim 20; SEQ ID NO 39253; 103pp; English.
 XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 ftp://wipo.int/pub/published_pct_sequences

XX SQ Sequence 435 AA;
 Query Match 7.7%; Score 158; DB 4;
 Best Local Similarity 33.9%; Pred. No. 5.5e-06;
 Matches 39; Conservative 17; Mismatches 59; Indels 0; Gaps 0;
 QY 48 MKAAALSVREDALGNITIGRREGTDPELPATAVGSHFDSVRNGGMFDGTTAGVVCALLEAARY 107
 Db 12 MAASGLETRFDEVGNLYGRLNTEYPOQEVVLSGSHTDVNGGNLDQFGALAALWD 71
 QY 108 MLENGYVNRHPPFEFIAIVEEGARFSSGMLGGRAIAGLVADRELDSLVDDEDGVSV 162
 Db 72 LKTQYGAFLRTVEVAMAEEEGSRFPYVFWGSKNIFGLANPDDVFNICDAKGNSI 126

RESULT 15
 ABG25030 ID ABG25030 standard; protein; 502 AA.
 XX AC ABG25030;
 XX DT 18-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #25021.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder.
 OS Homo sapiens.
 XX PN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 11-OCT-2001.
 XX PR

DR N-PSDB; AAS59520.
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX Example 1; SEQ ID NO 4651; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by *P. acnes*. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis. *P. acnes* is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of *P. acnes* in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for *P. acnes* proteins. These antibodies can be used to downregulate expression and activity of *P. acnes* polypeptides and therefore treat *P. acnes* infections. The antibodies may also be used as diagnostic agents for determining *P. acnes* presence, for example, by enzyme linked immunosorbent assay (ELISA). Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 403 AA;

Query Match	21.7%;	Score 444;	DB 4;	Length 403;
Best Local Similarity	32.8%;	Pred. No. 4.1e-33;		
Matches 135;	Conservative 50;	Mismatches 198;	Indels 28;	Gaps 10;

QY 8 AARIEKEIRELRSRFAEAGPGVTRLTYTEHAAARETLIAAMKAALSVREDALGNIGRR 67
 DB 11 AASLLAEIAAIGR--NEDGSYSRFCLREEVALREWFKAKATEGLAIVTDANANIWAWW 68

QY 68 EGTDPELPAIVAVGSHFDSDVRNGMFDGTAGVYCALEAARVMLENGYVNRRHFFEFIAVEE 127
 DB 69 GTPGPD--AVLTGSHLDSPVGGAYDGPLGVVSSLVAVAKMQEAGIMPAPKFAVVAMADE 126

QY 128 EGARFSSGMLGGRAIAGLYADRELDLVLDEDGVSVRQAATAFGQLKPGEELQAAARSAAADLR 187
 DB 127 EGARFGMPCLGSRSLASGKLSSKAEAHTLWARDGQSLPDAWREAGLNP-DLMGPDDVLLAES 185

QY 188 AFIELHIBEQGPILQEQQTEIGVVTISVGVRALRVAVKGRSDHAGTTPMHLRQDALVPAA- 246
 DB 186 CFVELHVEQGRGLADVGHPVAIATAVRHGRWRAEFTGQGNHAGTTLIPDRHDPPVVMAQ 245

QY 247 --LMVREVNRFVNNEIADGTVATVGHLTVAPEGGGNQVPGEVDFTLDLRSPEESLRLVLDI 304
 DB 246 TILAAREAE---REGCVTTGRINVEPGGTNVIASAATMWLDCRA--EES----QT 293

QY 305 ISVMVGEVASQAGVAADVDEFNLSPVQLAPTMVD--AVREAMASALQFTHRDI--SSGA 359
 DB 294 VTKVVEDIOQQASETAADKO---GCQVVWVTRESWTDTAFLDLRARMQLQILGEDIPQLSTGA 350

XX SQ Sequence 403 AA;

Query Match	21.7%;	Score 444;	DB 6;	Length 403;
Best Local Similarity	32.8%;	Pred. No. 4.1e-33;		
Matches 135;	Conservative 50;	Mismatches 198;	Indels 28;	Gaps 10;

QY 8 AARIEKEIRELRSRFAEAGPGVTRLTYTEHAAARETLIAAMKAALSVREDALGNIGRR 67
 DB 11 AASLLAEIAAIGR--NEDGSYSRFCLREEVALREWFKAKATEGLAIVTDANANIWAWW 68

QY 68 EGTDPELPAIVAVGSHFDSDVRNGMFDGTAGVYCALEAARVMLENGYVNRRHFFEFIAVEE 127
 DB 69 GTPGPD--AVLTGSHLDSPVGGAYDGPLGVVSSLVAVAKMQEAGIMPAPKFAVVAMADE 126

QY 128 EGARFSSGMLGGRAIAGLYADRELDLVLDEDGVSVRQAATAFGQLKPGEELQAAARSAAADLR 187

carbohydrase, lipase, isomerase, oxidase, reductase, transferase, kinase or phosphatase (all claimed). Also claimed is a method for detecting a Gram positive microorganism MP using a probe comprising all or part of the YurH DNA

CC
CC
CC
XX
SQ Sequence 412 AA;

Db	QY	Query Match	Score	DB	Length
28	APGLQNQIDEDELSSFSADAPSPTSVTRVLTYTDKDVSAARRYVXNLMALAGLTREDAVGNIFGK	29.4%	2	412;	412;
67	REGTDPPELPATAVGSHFDSVRNGGMFDGTAGVVCALEAARVMLENGYVNRHPFFEFIAVE	Best Local Similarity	35.2%	2;	
88	WDGLEPNLPAVATGSHIDAIPIYSCKYDGVVGVLGAIEAINVLKRSGFKPRSLEIILFTS	Pred. No.	4.1e-48;		
Db	127	EGARFSSGMLGGRAIAGL--VADRELDSSLVEDGVSVRQAAATAFGL--KPGELQAAAR	Mismatches	3;	Gaps 2;
QY	148	EETPRFGISCLGSRLLLASKELAALKTIVVDGONVSFIEARSAGYAEDKDDLSSVFL	Matches	179;	Indels 3;
Db	182	SAADLRAFIELHIEQGPILEQEQIEIGVVTISIVGRALRVAVKGRSDHAAGTPMHLRQDA	137;	Conservative	70;
QY	208	KKGSYPAFILELHIEQGPILEDEGLDIGVTAIAAPASLKVEPEGNGGHAGAVLMPYRNDA	267		
Db	242	LVAALMVREVNRNFVNEIAD-GTVATVGHLTVAAPGGNNQVPGEVDFTLDRSPHEESLRV	241		
QY	268	GLAAELALAAVEKHVLESESIDTVGTGILELHPGAINSIPSKSHLEIDTRDIDEARRNT	300		
Db	301	LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMV-DAVREAASALQFTHRDIISSGA	359		
QY	328	VKKIQESANTIARKRKVKLSEEFKIVNQDPPALSQDKLVTKKMAEAATENLNSHKMMISRA	387		
Db	360	GHDSMFIAQVTDVGVMVFVPSRAGRSHVPEEWTFDDDLRKGTTEVV 403	201		
QY	388	YHDLSLFMARIISPAGMIFIPCYKGYSHKPPEEYSSPEDMANGVVKVL	431		
Db	127	EGARFSSGMLGGRAIAGL--VADRELDSSLVEDGVSVRQAAATAFGL--KPGELQAAAR	181		
QY	148	EETPRFGISCLGSRLLLASKELAALKTIVVDGONVSFIEARSAGYAEDKDDLSSVFL	207		
Db	182	SAADLRAFIELHIEQGPILEQEQIEIGVVTISIVGRALRVAVKGRSDHAAGTPMHLRQDA	241		
QY	208	KKGSYPAFILELHIEQGPILEDEGLDIGVTAIAAPASLKVEPEGNGGHAGAVLMPYRNDA	267		
Db	242	LVAALMVREVNRNFVNEIAD-GTVATVGHLTVAAPGGNNQVPGEVDFTLDRSPHEESLRV	300		
QY	268	GLAAELALAAVEKHVLESESIDTVGTGILELHPGAINSIPSKSHLEIDTRDIDEARRNT	327		
Db	301	LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMV-DAVREAASALQFTHRDIISSGA	359		
QY	328	VKKIQESANTIARKRKVKLSEEFKIVNQDPPALSQDKLVTKKMAEAATENLNSHKMMISRA	387		
Db	360	GHDSMFIAQVTDVGVMVFVPSRAGRSHVPEEWTFDDDLRKGTTEVV 403	201		
QY	388	YHDLSLFMARIISPAGMIFIPCYKGYSHKPPEEYSSPEDMANGVVKVL	431		
Db	127	EGARFSSGMLGGRAIAGL--VADRELDSSLVEDGVSVRQAAATAFGL--KPGELQAAAR	181		
QY	148	EETPRFGISCLGSRLLLASKELAALKTIVVDGONVSFIEARSAGYAEDKDDLSSVFL	207		
Db	182	SAADLRAFIELHIEQGPILEQEQIEIGVVTISIVGRALRVAVKGRSDHAAGTPMHLRQDA	241		
QY	208	KKGSYPAFILELHIEQGPILEDEGLDIGVTAIAAPASLKVEPEGNGGHAGAVLMPYRNDA	267		
Db	242	LVAALMVREVNRNFVNEIAD-GTVATVGHLTVAAPGGNNQVPGEVDFTLDRSPHEESLRV	300		
QY	268	GLAAELALAAVEKHVLESESIDTVGTGILELHPGAINSIPSKSHLEIDTRDIDEARRNT	327		
Db	301	LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMV-DAVREAASALQFTHRDIISSGA	359		
QY	328	VKKIQESANTIARKRKVKLSEEFKIVNQDPPALSQDKLVTKKMAEAATENLNSHKMMISRA	387		
Db	360	GHDSMFIAQVTDVGVMVFVPSRAGRSHVPEEWTFDDDLRKGTTEVV 403	201		
QY	388	YHDLSLFMARIISPAGMIFIPCYKGYSHKPPEEYSSPEDMANGVVKVL	431		
Db	127	EGARFSSGMLGGRAIAGL--VADRELDSSLVEDGVSVRQAAATAFGL--KPGELQAAAR	181		
QY	148	EETPRFGISCLGSRLLLASKELAALKTIVVDGONVSFIEARSAGYAEDKDDLSSVFL	207		
Db	182	SAADLRAFIELHIEQGPILEQEQIEIGVVTISIVGRALRVAVKGRSDHAAGTPMHLRQDA	241		
QY	208	KKGSYPAFILELHIEQGPILEDEGLDIGVTAIAAPASLKVEPEGNGGHAGAVLMPYRNDA	267		
Db	242	LVAALMVREVNRNFVNEIAD-GTVATVGHLTVAAPGGNNQVPGEVDFTLDRSPHEESLRV	300		
QY	268	GLAAELALAAVEKHVLESESIDTVGTGILELHPGAINSIPSKSHLEIDTRDIDEARRNT	327		
Db	301	LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMV-DAVREAASALQFTHRDIISSGA	359		
QY	328	VKKIQESANTIARKRKVKLSEEFKIVNQDPPALSQDKLVTKKMAEAATENLNSHKMMISRA	387		
Db	360	GHDSMFIAQVTDVGVMVFVPSRAGRSHVPEEWTFDDDLRKGTTEVV 403	201		
QY	388	YHDLSLFMARIISPAGMIFIPCYKGYSHKPPEEYSSPEDMANGVVKVL	431		
Db	127	EGARFSSGMLGGRAIAGL--VADRELDSSLVEDGVSVRQAAATAFGL--KPGELQAAAR	181		
QY	148	EETPRFGISCLGSRLLLASKELAALKTIVVDGONVSFIEARSAGYAEDKDDLSSVFL	207		
Db	182	SAADLRAFIELHIEQGPILEQEQIEIGVVTISIVGRALRVAVKGRSDHAAGTPMHLRQDA	241		
QY	208	KKGSYPAFILELHIEQGPILEDEGLDIGVTAIAAPASLKVEPEGNGGHAGAVLMPYRNDA	267		
Db	242	LVAALMVREVNRNFVNEIAD-GTVATVGHLTVAAPGGNNQVPGEVDFTLDRSPHEESLRV	300		
QY	268	GLAAELALAAVEKHVLESESIDTVGTGILELHPGAINSIPSKSHLEIDTRDIDEARRNT	327		
Db	301	LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMV-DAVREAASALQFTHRDIISSGA	359		
QY	328	VKKIQESANTIARKRKVKLSEEFKIVNQDPPALSQDKLVTKKMAEAATENLNSHKMMISRA	387		
Db	360	GHDSMFIAQVTDVGVMVFVPSRAGRSHVPEEWTFDDDLRKGTTEVV 403	201		
QY	388	YHDLSLFMARIISPAGMIFIPCYKGYSHKPPEEYSSPEDMANGVVKVL	431		
Db	127	EGARFSSGMLGGRAIAGL--VADRELDSSLVEDGVSVRQAAATAFGL--KPGELQAAAR	181		
QY	148	EETPRFGISCLGSRLLLASKELAALKTIVVDGONVSFIEARSAGYAEDKDDLSSVFL	207		
Db	182	SAADLRAFIELHIEQGPILEQEQIEIGVVTISIVGRALRVAVKGRSDHAAGTPMHLRQDA	241		
QY	208	KKGSYPAFILELHIEQGPILEDEGLDIGVTAIAAPASLKVEPEGNGGHAGAVLMPYRNDA	267		
Db	242	LVAALMVREVNRNFVNEIAD-GTVATVGHLTVAAPGGNNQVPGEVDFTLDRSPHEESLRV	300		
QY	268	GLAAELALAAVEKHVLESESIDTVGTGILELHPGAINSIPSKSHLEIDTRDIDEARRNT	327		
Db	301	LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMV-DAVREAASALQFTHRDIISSGA	359		
QY	328	VKKIQESANTIARKRKVKLSEEFKIVNQDPPALSQDKLVTKKMAEAATENLNSHKMMISRA	387		
Db	360	GHDSMFIAQVTDVGVMVFVPSRAGRSHVPEEWTFDDDLRKGTTEVV 403	201		
QY	388	YHDLSLFMARIISPAGMIFIPCYKGYSHKPPEEYSSPEDMANGVVKVL	431		
Db	127	EGARFSSGMLGGRAIAGL--VADRELDSSLVEDGVSVRQAAATAFGL--KPGELQAAAR	181		
QY	148	EETPRFGISCLGSRLLLASKELAALKTIVVDGONVSFIEARSAGYAEDKDDLSSVFL	207		
Db	182	SAADLRAFIELHIEQGPILEQEQIEIGVVTISIVGRALRVAVKGRSDHAAGTPMHLRQDA	241		
QY	208	KKGSYPAFILELHIEQGPILEDEGLDIGVTAIAAPASLKVEPEGNGGHAGAVLMPYRNDA	267		
Db	242	LVAALMVREVNRNFVNEIAD-GTVATVGHLTVAAPGGNNQVPGEVDFTLDRSPHEESLRV	300		
QY	268	GLAAELALAAVEKHVLESESIDTVGTGILELHPGAINSIPSKSHLEIDTRDIDEARRNT	327		
Db	301	LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMV-DAVREAASALQFTHRDIISSGA	359		
QY	328	VKKIQESANTIARKRKVKLSEEFKIVNQDPPALSQDKLVTKKMAEAATENLNSHKMMISRA	387		
Db	360	GHDSMFIAQVTDVGVMVFVPSRAGRSHVPEEWTFDDDLRKGTTEVV 403	201		
QY	388	YHDLSLFMARIISPAGMIFIPCYKGYSHKPPEEYSSPEDMANGVVKVL	431		
Db	127	EGARFSSGMLGGRAIAGL--VADRELDSSLVEDGVSVRQAAATAFGL--KPGELQAAAR	181		
QY	148	EETPRFGISCLGSRLLLASKELAALKTIVVDGONVSFIEARSAGYAEDKDDLSSVFL	207		
Db	182	SAADLRAFIELHIEQGPILEQEQIEIGVVTISIVGRALRVAVKGRSDHAAGTPMHLRQDA	241		
QY	208	KKGSYPAFILELHIEQGPILEDEGLDIGVTAIAAPASLKVEPEGNGGHAGAVLMPYRNDA	267		
Db	242	LVAALMVREVNRNFVNEIAD-GTVATVGHLTVAAPGGNNQVPGEVDFTLDRSPHEESLRV	300		
QY	268	GLAAELALAAVEKHVLESESIDTVGTGILELHPGAINSIPSKSHLEIDTRDIDEARRNT	327		
Db	301	LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMV-DAVREAASALQFTHRDIISSGA	359		
QY	328	VKKIQESANTIARKRKVKLSEEFKIVNQDPPALSQDKLVTKKMAEAATENLNSHKMMISRA	387		
Db	360	GHDSMFIAQVTDVGVMVFVPSRAGRSHVPEEWTFDDDLRKGTTEVV 403	201		
QY	388	YHDLSLFMARIISPAGMIFIPCYKGYSHKPPEEYSSPEDMANGVVKVL	431		
Db	127	EGARFSSGMLGGRAIAGL--VADRELDSSLVEDGVSVRQAAATAFGL--KPGELQAAAR	181		
QY	148	EETPRFGISCLGSRLLLASKELAALKTIVVDGONVSFIEARSAGYAEDKDDLSSVFL	207		
Db	182	SAADLRAFIELHIEQGPILEQEQIEIGVVTISIVGRALRVAVKGRSDHAAGTPMHLRQDA	241		
QY	208	KKGSYPAFILELHIEQGPILEDEGLDIGVTAIAAPASLKVEPEGNGGHAGAVLMPYRNDA	267		
Db	242	LVAALMVREVNRNFVNEIAD-GTVATVGHLTVAAPGGNNQVPGEVDFTLDRSPHEESLRV	300		
QY	268	GLAAELALAAVEKHVLESESIDTVGTGILELHPGAINSIPSKSHLEIDTRDIDEARRNT	327		
Db	301	LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMV-DAVREAASALQFTHRDIISSGA	359		
QY	328	VKKIQESANTIARKRKVKLSEEFKIVNQDPPALSQDKLVTKKMAEAATENLNSHKMMISRA	387		
Db	360	GHDSMFIAQVTDVGVMVFVPSRAGRSHVPEEWTFDDDLRKGTTEVV 403	201		
QY	388	YHDLSLFMARIISPAGMIFIPCYKGYSHKPPEEYSSPEDMANGVVKVL	431		
Db	127	EGARFSSGMLGGRAIAGL--VADRELDSSLVEDGVSVRQAAATAFGL--KPGELQAAAR	181		
QY	148	EETPRFGISCLGSRLLLASKELAALKTIVVDGONVSFIEARSAGYAEDKDDLSSVFL	207		
Db	182	SAADLRAFIELHIEQGPILEQEQIEIGVVTISIVGRALRVAVKGRSDHAAGTPMHLRQDA	241		
QY	208	KKGSYPAFILELHIEQGPILEDEGLDIGVTAIAAPASLKVEPEGNGGHAGAVLMPYRNDA	267		
Db	242	LVAALMVREVNRNFVNEIAD-GTVATVGHLTVAAPGGNNQVPGEVDFTLDRSPHEESLRV	300		
QY	268	GLAAELALAAVEKHVLESESIDTVGTGILELHPGAINSIPSKSHLEIDTRDIDEARRNT	327		
Db	301	LIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMV-DAVREAASALQFTHRDIISSGA	359		
QY	328	VKKIQESANTIARKRKVKLSEEFKIVNQDPPALSQDKLVTKKMAEAATENLNSHKMMISRA	387		
Db	360	GHDSMFIAQVTDVGVMVFVPSRAGRSHVPEEWTFDDDLRKGTTEVV 403	201		
QY	388	YHDLSLFMARIISPAGMIFIPCYKGYSHKPPEEYSSPEDMANGVVKVL	431		
Db	127	EGARFSSGMLGGRAIAGL--VADRELDSSLVEDGVSVRQAAATAFGL--KPGELQAAAR	181		
QY	148	EETPRFGISCLGSRLLLASKELAALKTIVVDGONVSFIEARSAGYAEDKDDLSSVFL	207		
Db	182	SAADLRAFIELHIEQGPILEQEQIEIGVVTISIVGRALRVAVKGRSDHAAGTPMHLRQDA	241		
QY	208	KKGSYPAFILELHIEQGPILEDEGLDIGVTAIAAPASLKVEPEGNGGHAGAVLMPYRNDA	267		
Db	242	LVAALMVREVNRNFVNEIAD-GTVATVGHLTVAAPGGNNQVPGEVDFTLDRSPHEESLRV	300		
QY	268	GLAAELALAAVEKHVLESESIDTVGTGILELHPGAINSIPSKSHLEIDTRDIDEARRNT</td			

Qy	369 VTDVGMOVFPVPSRAGRSHVPEEWTDFDLRKGTTEVVLRYMK 408	Db	300 RRNKIIKEKIEEKIKLIVSNTRGLEYQIEKNMMAAVPVVKCSENLINSLKQSCKBLEIDAPIIV 359
Db	366 LTEVGLIFVPSHNGISHAPEEWTDKLQKGIEVVVKTVK 405	Qy	357 SGAGHDMSMFIAQVTDVGMVFVPSRAGRSHVPEEWTDFDLRKGTTEV 403
		Db	360 SGAGHDAMFLAEITEIGMVFVRCRNGISHSPKEWAEIDDILTGTVK 406
RESULT 5		RESULT 6	
AAR15484	ID AAR15484 standard; protein: 414 AA.	AAR25693	
XX	AC AAR15484;	ID AAR25693	standard; protein: 409 AA.
XX	DT 25-MAR-2003 (revised)	XX	
DT 17-MAR-1992 (first entry)	AC AAR25693;	XX	
DE Protein encoded by ORF 4 contained in PHPB12.	DT 28-JAN-1993 (first entry)	XX	
XX	DE Heat resistant carbamylase.	XX	
KW Hydrolase; 5-substituted hydantoin; N-carbamyl amino acid.	KW KW Temp; stable; N-carbamyl-L amino acid.	XX	
XX	OS Pseudomonas NS 671 (FBRM P-9543).	XX	
OS	PN JP03251176-A.	OS Escherichia coli.	
XX	PD 08-NOV-1991.	PN JP04183391-A.	
XX	PP 29-MAY-1990;	XX 30-JUN-1992.	
XX	PR 10-JAN-1990;	PP 15-NOV-1990;	90JP-00307221.
PR 90JP-00002786.	XX 15-NOV-1990;	PR 15-NOV-1990;	90JP-00307221.
XX	PA (NIPS) NIPPON SODA CO.	XX PA (NIPS) NIPPON SODA CO.	
XX	DR WPI; 1991-373414/51.	XX DR WPI; 1992-265588/32.	
DR N-PSDB; AAQ15312.	DR P-PSDB; AAR25693.	XX DR P-PSDB; AAR25693.	
PT Enzyme for transforming 5-substd. hydantoin - in which enzymatic protein	PT Heat resistant carbamylase - is used to produce L-aminoacid in high	XX PT PT purity by reacting enzyme protein with N-carbamyl-L-aminoacid.	
PT has specified aminoacid sequence.	XX PS Claim 1; Fig 3; 8pp; Japanese.	XX PS	
XX	PS XX	PS	
CC The sequence is one of three proteins encoded by three open reading	CC The heat resistant carbamylase has an optimum temp. ca. 60 degrees C and	XX CC	
CC frames: 2, 3, and 4, which together encode an enzyme which has the ability	CC an optimum pH ca. 8.5 and can be stably produced by bacteria. The enzyme	CC	
CC to transform 5-substd. hydantoin into N-carbamyl amino acids. The DNA	CC allows efficient prodn. of a high concn. of N-carbamyl-L-amino acids	CC	
CC segment contg. the three ORFs was obtnd. from Pseudomonas NS 671 and	CC which may be hydrolysed to the relavent L- amino acids in the presence of	CC	
CC ligated into pUC18 to form the plasmid PHPB12. See also AAR15482 and	CC the heat resistant carbamylase	CC	
CC AAR15483. (Updated on 25-MAR-2003 to correct PA field.)	XX SQ Sequence 414 AA;	XX SQ Sequence 409 AA;	
SQ	XX	XX	
Query Match 35.5%; Score 727; DB 2; Length 414;	Query Match 35.5%; Score 726.5%; DB 2; Length 409;	XX	
Best Local Similarity 37.3%; Pred. No. 6.8e-60;	Best Local Similarity 38.8%; Pred. No. 7.4e-60;	XX	
Matches 152; Conservative 82; Mismatches 165; Indels 8; Gaps 5;	Matches 156; Conservative 72; Mismatches 163; Indels 11; Gaps 4;	XX	
Qy 1 MTLOKAQARIEKEIRELSRF-SAEGPGVTRLTYTPHEAAARETLLIAAMKAAALSVERDA 59	Qy 7 QAARIEKEIRELSRFSAE-GPGVTRLTYTPHEAAARETLLIAAMKAAALSVERDALGNIG 65	XX	
Db 4 VRISK---ERLRTHIQLGEIGKTKDKGVQRLLSKEDREATLVSEWRAGLTVTVDH 60	Db 3 QGERLWQLMELGEVQKOPSGGVTRLSFTAERRAKDLYASYMREAGLFVYEDAAGNLIG 62	XX	
Qy 60 LGNIIGRREGTDPPELPALIAVGSHFDSVRNGGMFDGTTAGGVCALEAARVMLENGYVNRRHPP 119	Qy 66 RREGTDPPELPALIAVGSHFDSVRNGGMFDGTTAGGVCALEAARVMLENGYVNRRHPPFIAIV 125	XX	
Db 61 FGNLIGRKEGETPSSLPSVMIGSHIDSVRNGKFDGVIGVLAGIEIVHAISEANVVEHSI 120	Db 63 RKEGTNPDPATVVLVGSHLDSVYNNGGCFDGPLGVLAGVEVQTNEHGVVTHPIEVVAFT 122	XX	
Qy 120 EFAIAIVEEGARFSSGMLGGRAIAGLVADRELSDLVDEDGVSVRQA-ATAFGLKPKGELQ 177	Qy 126 EEEGARFSSGMLGGRAIAGLVADRELSDLVDEDGVSVRQAATAFGLKPKGELQAAARSAA 185	XX	
Db 121 EVVAFCEEESRFNDGLFGSRGMVGKVKPEDLQK-VDDNNVTRYEARLKTFGFGIDPDTF 179	Db 123 DEEEGARFRGMGSRAAMAGTLPPREAECR-DAEGISLQDPLDRQAAARKPGT 181	XX	
Qy 178 AAARSAAIDLRAFILELHIEQGFILEEQEIQIEIGVVTTSIVGVRALRVAVKGRSDHAGTTPMHL 237	Qy 186 LRAFIELHIEQGPILEQEQIEIGVVTTSIVGVRALRVAVKGRSDHAGTTPMHL 245	XX	
Db 180 QSIREIGDIKHYFEMHIEQGPYLEKNINYPIGIVSGVIAQPSWFKVRLVGEAGHAGTVPM 239	Db 182 VKAYVELHIEQGRVLEETGLPVGVITGIAGLIWVKFTIEGKAEHAGATPMSSLRDPMAA 241	XX	
Qy 238 RODALVPAALMVRENVNRF-VNEIADGTVATVGHLTVAPEGGNQVPGEVDFTLDLRSPHEE 296	Qy 246 ALMV---REVNRFVNEIADGTVATVGHLTVAPEGGNQVPGEVDFTLDLRSPHEE 301	XX	
Db 240 RKDPLVGAABEVIKEVETLCMNDPNAPTVGTGRIAAPPGGSNIIPESVEFTLDIRDIE 299	Db 242 AQIIIVIEEARR----TGTGTVGQHLVYPGGINVIPERFEVFLDLRDKAEVRDQV 296	XX	
Qy 297 SLRVLIDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMVDVREAAASALQFTHRDIS 356	Qy 302 IDRISVMVGEVASQAGVAADVDEFNLSPVQLAPTMVDVREAAASALQFTHRDIS 356	XX	

XX	Suzuki S, Onishi N, Yokozeiki K;	PD	18-OCT-2001.
XX	WPI; 2002-510588/55.	XX	11-APR-2001; 2001WO-FR001118.
DR	N-PSDB; ABN86382, ABN86383.	XX	11-APR-2000; 2000FR-00004629.
XX	New 5-substituted hydantoin racemase, useful in production of optically active amino acids, comprises high working temperature, from microbacterium species.	XX	(INST PASTEUR.
PT	Page 27-28; 40pp; English.	XX	Buchrieser C, Frangeul L, Couve E, Rusniok C, Fsihi H, Dehoux P, Cossart P;
PS	The invention relates to a 5-substituted hydantoin racemase (HRase) that is derived from a Microbacterium by culture, disruption or lysis, and purification. HRase has a high working temperature of 5-60 plusOC and is used to racemize optically active 5-substituted hydantoins for subsequent enzymatic conversion to N-carbamoyl-amino acids and then optically active amino acids (useful in pharmaceuticals, the chemical industry and as food additives). The present sequence represents the M. liquefaciens N-carbamoyl-amino acid-L-hydrolase (Chase) enzyme.	XX	Dussurget O, Chetouani F, Nedjari H, Glaser P, Kunst F, Vazquez-Boland JA; Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A; Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L; Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N; Madueno E, De Fablos B, Weiland J, Kaerst U, Entian K, Hauf J; Rose M, Voss H;
PT	Sequence 412 AA;	XX	WPI; 2002-010914/01.
PS	Genomic sequence for Listeria monocytogenes, useful e.g. for treatment and prevention of Listeria and related bacterial infections, and related polypeptides.	XX	DR
PT	Claim 6; SEQ ID NO 80; 192pp; French.	XX	PT
PT	The present invention relates to the genome sequence of Listeria monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of it are useful for selecting probes and primers for detecting genes in L. monocytogenes and related organisms, and for studying genetic polymorphisms and other genomes. The present sequence is a protein encoded by the genome sequence of the present invention. Proteins expressed from the genome sequence are useful for raising specific antibodies, identification of L. monocytogenes and related organisms, and for biosynthesis and biodegradation, especially biosynthesis of Vitamin B12. The genome sequence and proteins encoded by it are also useful for selecting compounds that regulate gene expression and cell replication and modulate L. monocytogenes-related diseases. In addition, the genome sequence and proteins encoded by it are useful in pharmaceutical and vaccines compositions for the treatment or prevention of infections by L. monocytogenes and related organisms. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences	XX	PT
PT	Sequence 423 AA;	XX	SQ
Query	Match 83.4%; Score 1709; DB 5; Length 412;	PS	Query Match 41.9%; Score 858; DB 5; Length 423;
Best Local Similarity 81.1%; Pred. No. 7e-153;	Best Local Similarity 44.2%; Pred. No. 2.8e-72;	Matches 334; Conservative 38; Mismatches 40; Indels 0; Gaps 0;	Best Local Similarity 44.2%; Pred. No. 2.8e-72;
Matches 334; Conservative 38; Mismatches 40; Indels 0; Gaps 0;	Matches 177; Conservative 68; Mismatches 153; Indels 2; Gaps 2;	CC	Matches 177; Conservative 68; Mismatches 153; Indels 2; Gaps 2;
Db	1 MTLOKAQAAARIEKEIRELTSRFSAEAGPGVTRLTYTPHEAAARETLLIAAMKAALSVREDAL 60	CC	CC
Db	1 VTLQQARADRIEEELWLTSRFSVEGPVGTRLTYTPHEAAAREVVAAMQRGTLGSVHEDAL 60	CC	CC
QY	61 GNIIGREGTDPPELPAIAVGSHFDSVRNGGMFDGTAGVVCALEARVMLENGYVNRRHPE 120	CC	CC
Db	61 GNIIGREGSDPDLPAIAFGSHFDSVRNGGMFDGTAGVVCALEARVLQESGYVNRRHPE 120	CC	CC
QY	121 FIAIVEEGARFSGMLGGRAIAGLVADRELDSLVDDEDGVSVRQAATAFGLKPGELQAA 180	CC	CC
Db	121 VIAIVEEGTRFSGMLGGRAIAGLVSDADDLTVDEDGVTVREATAFGLEPGELRTAA 180	CC	CC
QY	181 RSAADLRAFIELHIEQPILEEQEIQIEIGVVTSIVGVRLAVKGRSDDHAGTTPMHILRQD 240	CC	CC
Db	181 RTRDDLRAFIELHIEQPILEQEKEVIGVTVAFRITVEGRSDDHAGTTPMHILRQD 240	CC	CC
QY	241 ALVPAAALMVRREVNRFVNIEIDAGTVATVGHLTIVAPGGGNQVPGEVDFTLDLRSPEEESLRV 300	XX	CC
Db	241 ALVPAAALMVRREVNRFVNIEIDAGTVATVGHLTIVAPGGGNQVPGEVDFTLDLRSPEEESIRL 300	XX	CC
QY	301 LIDRISWMGEVASQAGVAADVDEFNLSPVQLAPTMVDAVREAAASALQFTHRDISSGAG 360	QY	10 RIEKEIRELSRFSA-EGPGVTRLTYTPHEAAARETLLIAAMKAALSVREDALGNIGRRE 68
Db	301 LVDRIEAMVAEAAAAGVEAVNGFFALSPVGILSPVVDRVRDAASELGFTHRDITSGAG 360	Db	7 RIKKHIEKLDTFTATPGQGTTRLTYSKEDLDARNYLQKQEMAKVGVLTVSEDAIGNIYGRLE 66
QY	361 HDMSMFIAQVTDVGMVFVPSRAGRSHVPEWTDDDLRKGTTEVYLRVMKALDR 412	QY	69 GTDPELPALAVGSHFDSVRNGGMFDGTAGVVCALEARVMLENGYVNRRHPEFIAVEEE 128
Db	361 HDMSMFIAQVTDVGMVFVPSRAGRSHVPEWEWSDDFLRKGTDVVLHYVTAALDR 412	Db	67 GESPDLPAVIVGSHFDSVPNGGAFDGPAGVITGLEVASVFEQQQIKPHFPLETTIAMVEEE 126
QY	RESULT 4	QY	129 GARFSSGMGGRATAGLVADELDSLVDGEDGVSVRQAAATAFGLKEGELQAARSAADLRA 188
ID	ABB47375 standard; protein; 423 AA.	XX	127 GSREFAGLILASRAITGKVTTMHEMKDIDGITATDAMAKLGFNANRVRHAIRTKEVKAA 186
XX	ABB47375;	Db	189 FIELHIEQGPILEOEQIEIGVVTSTIVGVRALRVAVKGRSRDHAGTTPMHILRODALVPAALM 248
AC		Db	187 FIELHIEQGPILENASEDVALVTDVGLTQIKVTVKGQAGHGTTPMLDRKDVLVSAVEI 246
XX	05-FEB-2002 (first entry)	QY	249 VREVNRFVNEIADGTVATVGHLTIVAPGGNNQVPGEVDFTLDLRSPEEESLRVLDRISVM 308
DE	Listeria monocytogenes protein #79.	Db	247 LGQLPELAIQEGGGTVLTVGKLNVYPNGANVIPDKVIFTDIRAKDEIHVNQNTLAKTKEI 306
XX	Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation; vitamin B12; bacterial infection; disease.	QY	309 VGEVASQGVAADVDEFNLSPVQLAPTMVDAVREAASALQFTHRDISSGAGHDMSMFIAQ 368
OS	Listeria monocytogenes.	Db	307 1-QATEKNITCEIKDMIYQQPHTLSKEIHQALTESADOLGFKYRTMVSGAGHDAMIFAS 365
XX	WO200177335-A2.	PN	
XX		XX	

CC Isolation of microorganisms capable for stereospecific hydrolysis of
 CC hydantoins is necessary due to the increasing demand for optically pure
 CC amino acids. The cells are usually bacterial cells, e.g. from Escherichia
 CC coli, since they have high reproduction rates and easy growing
 CC conditions. Since all the enzymes are produced in one strain, there is a
 CC reduction of fermentation and purification costs
 XX

SQ Sequence 412 AA;

Query Match 100.0%; Score 2049; DB 4; Length 412;
 Best Local Similarity 100.0%; Pred. No. 4.5e-185;
 Matches 412; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MTLQKAQAARIEKEIRELRSRFSAEGPGVTRLTYTPEHAAARETLIAAMKAALSVREDAL 60
 1 MTLQKAQAARIEKEIRELRSRFSAEGPGVTRLTYTPEHAAARETLIAAMKAALSVREDAL 60
 Qy 61 GNIIGRREGTDPELPAIAGSHFDSVRNGGMFDGTAGVYCALEAARVMJENGYVNRRHPFE 120
 61 GNIIGRREGTDPELPAIAGSHFDSVRNGGMFDGTAGVYCALEAARVMJENGYVNRRHPFE 120
 Db 61 GNIIGRREGTDPELPAIAGSHFDSVRNGGMFDGTAGVYCALEAARVMJENGYVNRRHPFE 120

Qy 121 FIAIVEEGARFSSGMLGGRAIAGLVADRELDLSLVEDDGVSVRQAAATAFGLKPGELQAAA 180
 Db 121 FIAIVEEGARFSSGMLGGRAIAGLVADRELDLSLVEDDGVSVRQAAATAFGLKPGELQAAA 180
 181 RSAADLRAPIELHIEQGPILEQEQIEIGVVTSIVGVRALRVAVKGRSDHAGTTPMHLRQD 240
 Qy 181 RSAADLRAPIELHIEQGPILEQEQIEIGVVTSIVGVRALRVAVKGRSDHAGTTPMHLRQD 240
 Db 181 RSAADLRAPIELHIEQGPILEQEQIEIGVVTSIVGVRALRVAVKGRSDHAGTTPMHLRQD 240
 Qy 241 ALVPAAALM\REVRNFVNNEIADGTATVGHLTVAPEGGNQVPQLAPTMVDLRLSPHEESLRV 300
 Db 241 ALVPAAALM\REVRNFVNNEIADGTATVGHLTVAPEGGNQVPQLAPTMVDLRLSPHEEESLRV 300
 Qy 241 ALVPAAALM\REVRNFVNNEIADGTATVGHLTVAPEGGNQVPQLAPTMVDLRLSPHEEESLRV 300
 Db 241 ALVPAAALM\REVRNFVNNEIADGTATVGHLTVAPEGGNQVPQLAPTMVDLRLSPHEEESLRV 300
 Qy 301 LIDRISVMGEVASQAGVAADVDEFNLSPVQLAPTMVDAREAAASALQFTHRDISSGAG 360
 Db 301 LIDRISVMGEVASQAGVAADVDEFNLSPVQLAPTMVDAREAAASALQFTHRDISSGAG 360
 Qy 361 HDSMFIQAQTVDGMVFVPSRAGRSHVPEWTFDDLRKGTEVVLRYMKALDR 412
 Db 361 HDSMFIQAQTVDGMVFVPSRAGRSHVPEWTFDDLRKGTEVVLRYMKALDR 412

RESULT 2
 ABB99394 ID ABB99394 standard; protein; 412 AA.
 XX AC ABB99394;
 XX DT 29-JAN-2003 (first entry)

XX DE Amino acid sequence of N-carbamyl-L-amino acid hydrolase enzyme.
 XX KW Hydantoinase; enzyme; N-carbamyl-L-amino acid hydrolase;
 XX optically-active amino acid.
 OS Microbacterium liquefaciens.
 PN WO200272841-A1.
 XX PR 08-MAR-2002; 2002WO-JP002173.
 XX PD 19-SEP-2002.
 XX PF 08-MAR-2002; 2002WO-JP002173.

XX PR 08-MAR-2001; 2001JP-00065814.
 PR 27-SEP-2001; 2001JP-00298619.
 XX PA (AJIN) AJINOMOTO CO INC.
 XX PR 13-SEP-2000; 2000JP-00278571.
 PR 08-MAR-2001; 2001JP-00065815.
 DR WPI; 2002-691806/74.
 DR N-PSDB; ABV72490.
 XX PA (AJIN) AJINOMOTO CO INC.

PT Microbacterium-originated DNAs encoding hydantoinase and N-carbamyl-L-amino-acid hydrolase and transformant cells, useful for production of proteins and optically-active amino acids, applicable in synthesis of e.g. pharmaceuticals.

XX PS Claim 14; Page 67-70; 83pp; Japanese.

XX SQ Sequence 412 AA;
 CC The present sequence represents N-carbamyl-L-amino acid hydrolase enzyme.
 CC The specification also describes a hydantoinase enzyme. The enzymes are
 CC used for production of proteins and optically-active amino acids, which
 CC are applicable in synthesis of pharmaceuticals, chemicals or food
 additives

XX SQ Sequence 412 AA;
 PT Query Match 83.4%; Score 1709; DB 5; Length 412;
 PT Best Local Similarity 81.1%; Pred. No. 7e-153;
 PT Matches 334; Conservative 38; Mismatches 40; Indels 0; Gaps 0;
 PT XX
 PT Query 1 MTLQKAQAARIEKEIRELRSRFSAEGPGVTRLTYTPEHAAARETLIAAMKAALSVREDAL 60
 PT Db 1 VTLQQARADRIEEELWTLSRFSVEGPGVTRLTYTPEHAAAREVAAAMORTGLSVHEDAL 60
 PT Qy 61 GNIIGRREGTDPELPAIAGSHFDSVRNGGMFDGTAGVYCALEAARVMJENGYVNRRHPFE 120
 PT Db 61 GNIIGRREGSDPALPAIAGSHFDSVRNGGMFDGTAGVYCALEAARVMJENGYVNRRHPFE 120
 PT Qy 121 FIAIVEEGARFSSGMLGGRAIAGLVADRELDLSLVEDDGVSVRQAAATAFGLKPGELQAAA 180
 PT Db 121 FIAIVEEGARFSSGMLGGRAIAGLVADRELDLSLVEDDGVSVRQAAATAFGLKPGELQAAA 180
 PT Qy 181 RSAADLRAPIELHIEQGPILEQEQIEIGVVTSIVGVRALRVAVKGRSDHAGTTPMHLRQD 240
 PT Db 181 RSAADLRAPIELHIEQGPILEQEQIEIGVVTSIVGVRALRVAVKGRSDHAGTTPMHLRQD 240
 PT Qy 241 ALVPAAALM\REVRNFVNNEIADGTATVGHLTVAPEGGNQVPQLAPTMVDLRLSPHEEESLRV 300
 PT Db 241 ALVPAAALM\REVRNFVNNEIADGTATVGHLTVAPEGGNQVPQLAPTMVDLRLSPHEEESLRV 300
 PT Qy 301 LIDRISVMGEVASQAGVAADVDEFNLSPVQLAPTMVDAREAAASALQFTHRDISSGAG 360
 PT Db 301 LIDRISVMGEVASQAGVAADVDEFNLSPVQLAPTMVDAREAAASALQFTHRDISSGAG 360
 PT Qy 361 HDSMFIQAQTVDGMVFVPSRAGRSHVPEWTFDDLRKGTEVVLRYMKALDR 412
 PT Db 361 HDSMFIQAQTVDGMVFVPSRAGRSHVPEWTFDDLRKGTEVVLRYMKALDR 412
 PT Qy 361 HDSMFIQAQTVDGMVFVPSRAGRSHVPEWTFDDLRKGTEVVLRYMKALDR 412
 PT Db 361 HDSMFIQAQTVDGMVFVPSRAGRSHVPEWTFDDLRKGTEVVLRYMKALDR 412

XX RESULT 3
 XX ABB80945 ID ABB80945 standard; protein; 412 AA.
 XX AC ABB80945;
 XX DT 21-OCT-2002 (first entry)
 XX DE M. liquefaciens Chase enzyme.
 XX OS Microbacterium liquefaciens.
 XX KW Hydantoin racemase; HRase; bacterium; hydantoin; N-carbamoyl-amino acid;
 XX KW pharmaceutical; chemical; food additive; hydantoinase; Chase; enzyme.
 XX PD 20-MAR-2002.
 XX PF 09-AUG-2001; 2001EP-00119275.
 XX PN EP1188826-A2.
 XX PR 13-SEP-2000; 2000JP-00278571.
 PR 08-MAR-2001; 2001JP-00065815.
 XX PA (AJIN) AJINOMOTO CO INC.

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 3, 2004, 18:14:00 ; Search time 58 Seconds
 (without alignments)
 2007.062 Million cell updates/sec

Title: US-10-045-063-2

Perfect score: 2049

Sequence: 1 MTLQKAQAAARIEKEIRELSR.....FDQLRKGTTEVVLRVMKALDR 412

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%
 Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

	RESULT 1	
ID	AAU01557	
AAU01557 standard; protein; 412 AA.		
XX		
AC		
XX		
XX		
DT	18-JUL-2001 (first entry)	
XX		
DE	Arthrobacter aurescens carbamoylase (hyuC).	
XX		
KW	Hydantoinase; racemase; carbamoylase; whole cell catalyst; L-amino acid;	
KW	D-amino acid; 5-monosubstituted hydantoin; racemisation; hydrolysis;	
KW	bacterial cell; Escherichia coli; purification; fermentation; hyuC; reproduction rate.	
XX		
OS	Arthrobacter aurescens.	
XX		
PN	WO2001233582-A1.	
XX		
PD	05-APR-2001.	
XX		
PP	31-AUG-2000; 2000WO-EP008473.	
XX		
PR	28-SEP-1999; 99US-00407062.	
XX		
(DEGS)	DEGUSSA-HUELS AG.	
(UYST-)	UNIV STUTTGART.	
(HOFF)	ROCHE DIAGNOSTICS GMBH.	
XX		
XX	Altenbuchner J, Matthes R, Syldatk C, Wiese A, Wilms B.	
PI	Bommarius A, Tischer W;	
XX		
WPI; 2001-266169/27.		
DR	N-PSDB; AAS02276.	
XX		
New whole cell catalyst for degrading hydantoins into amino acids, comprises a hydantoinase, a racemase and a carbamoylase.		
PT		
Example; Page 34-35; 51pp; English.		

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	2049	100.0	412 4 AAU01557	Aau01557 Arthrobac
2	1709	83.4	412 5 ABB99394	Abb99394 Amino aci
3	1709	83.4	412 5 ABB80945	Abb80945 M. liquef
4	858	41.9	423 5 ABB47375	Abb47375 Listeria
5	727	35.5	414 2 AAR15484	Aar15484 Protein e
6	726.5	35.5	409 2 AAR25693	Aar25693 Heat resi
7	719.5	35.1	409 2 AAW03544	Aaw03544 Bacillus
8	636	31.0	441 5 ABB93697	Abb93697 Herbicida
9	602.5	29.4	412 2 AAY06396	Aay06396 Bacillus
10	444	21.7	403 4 AAU43456	Aau43456 Propionib
11	444	21.7	403 6 ABM39975	Abm39975 Propionib
12	395.5	19.3	435 4 AAB97165	Aab97165 Pseudomon
13	349.5	17.1	554 4 ABG28578	Abg28578 Novel hum
14	158	7.7	435 4 ABG08894	Abg08894 Novel hum
15	146	7.1	502 4 ABG25030	Abg25030 Novel hum
16	146	7.1	502 4 ABG25687	Abg25687 Novel hum
17	143.5	7.0	319 6 ABU19682	Protein e
18	143.5	7.0	465 6 ADA36028	Ada36028 Acinetoba
19	141	6.9	346 2 AAR67430	Aar67430 Thermosta
20	140.5	6.9	447 6 ADA35069	Ada35069 Acinetoba
21	139	6.8	349 4 AAB96717	Aab96717 Putative
22	138.5	6.8	418 6 ABU22434	Abu22434 Protein e
23	133	6.5	406 6 ADA34022	Ada34022 Acinetoba
24	130.5	6.4	415 6 ABM71060	Abm71060 Staphyloc
25	127.5	6.2	383 5 ABP65813	Abp65813 Bifidobac

The sequence represents an Arthrobacter aurescens carbamoylase (hyuC). A whole cell catalyst for the conversion of 5-monosubstituted hydantoins to L- or D-amino acids, comprises cloned genes encoding for a hydantoinase, a racemase and a carbamoylase. These genes are overexpressed in the cell according to their turnover rates. The bioconversion consists of chemical and/or enzymatic racemisation of 5-substituted hydantoins, ring opening hydrolysis achieved by a hydantoinase, and hydrolysis of the N-carbamoyl amino acid produced by hydantoinase to the amino acid by carbamoylase.

source	Query	Match	Score	DB	Length
	Best	Local	Similarity	13;	669;
	Matches	181;	Conservative	0;	3.4e-05;
Qy	82	GTTACCGGGCTGACACTCCAGAGCATGCCGCGGGAAACGGCTCATTGGGGCT	141		
Db	272	GTGACCCGCATTCTGTACAGTATGAATGATGTTAGCTCGCAGGTATGTTAAAAGCCTT	331		
Qy	142	ATGAAAGGGCCGGCTTGAGCGGTTCTGTGAGAACGGCACTCGGAAACATCATGGCGGAGCT	201		
Db	332	ATGGAGGGAGCTGGCTCTCGTGTCCGTGAGGATGCTGTAGGAATAATTGGTCGTTGG	391		
Qy	202	GAAGGCACTGATCCGGAGCTTCTGGATCGGGTGGTTCACACTTCGATTCTGTCCGA	261		
Db	392	GAGGAAGCAATCAGAGTTATCTGCTGTAGCAAATGGATCTCATATTGGCTATTCCA	451		
Qy	262	AACGGGGATGTTGATGGCACTGCAGGGTGGCTGGCTGGCTGGGGCTGGGGCTGGGGTG	321		
Db	452	TACTCAGGCAAATTGATGGAGTTGGCTGGGGTCTGGGTGCCATTGAAGCTATAATGTG	511		
Qy	322	ATGCTGGAGAACGGCTACGTGAAATGGCATCCATTGAGTTCACTGGCATGGGGAG	381		
Db	512	CTGAGAAGGTTGGCATTCAAAACAAAGAGATCACTGGGGTGAATTATGTTACCTCTGAG	571		
Qy	382	GAAGGGCCGGCTCAGCAGGGCATGTTGGGGCCATTGGCAGGGTTGGT	437		
Db	572	GAGCCTACACGGTTGGAAATTGGCTGGCTGGGGTGTATGGCAGGGAGTGT	627		

Search completed: May 4, 2004, 16:01:56
 Job time : 3701 secs

Query Match 6.2%; Score 76.6; DB 12; Length 514;
 Best Local Similarity 51.1%; Pred. No. 2.3e-05;
 Matches 179; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Qy 82 GTTACCCGGCTGACCTTACACTCCAGGCAATGCCCGGGAAACGCTCATGGGGCT 141
 Db 138 GTAACCAGGGTCTGTGATACTGACAAGGATGTTAGGTGCAAGGTATGTAACAT 197

Qy 142 ATGAAAGGGCCCTTGAGGTTAGAACATCATCGGAACGACTCGGAAACATCATTGGCAAGT 201
 Db 198 ATGGAACATTGCTGGCTATCTGTGAGAAGGGATGCTGGTCGTGG 257

Qy 202 GAAGGCAGTGTACCGAGCTTCCTGCGATCGGGTCAACACTCGATTCTGTCGA 261
 Db 258 GATGGCTTGAACCTGAGCTGCTGCAACAGGTTGCAACAGGTGCTGACATGGCTATACT 317

Qy 262 AACGGGGGATGTTGATGGCACTGGGACTTGAGGGTGGCTGGGGCTTGGGGT 321
 Db 318 TACTCGGAAAATATGATGGAGTTAGGTGCTATTGAAAGCCATCAGAGTC 377

Qy 322 ATGCTGGAGAACGGCTACGTGAATCGGATCCATTGAGTTCATCGGATCTGGGAGG 381
 Db 378 CTGAAAAGGTCTGGTTAAACCTAGAAGACCATTGGAAAGTCATTCATTCACTAGAA 437

Qy 382 GAAGGGGCCGCTTCAGGAGTGGCATGTTGGCCATTGGCAGG 431
 Db 438 GAACCAAACAGCTTGGAAATAGGTTGCTGGAAAGCCGCCTATGGCTGG 487

RESULT 14

LOCUS BI973078 565 bp mRNA linear EST 29-NOV-2001
 DEFINITION sag86f04.y2 Gm-c1084 Glycine max cDNA clone GENOME SYSTEMS CLONE
 ID: Gm-c1084-1040 5' similar to TR:085664 N-CARBAMYL-L-AMINO
 ACID AMIDOHYDROLASE. ; mRNA sequence.

ACCESSION BI973078
 VERSION BI973078.1 GI:16347483
 KEYWORDS EST.
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 565)

AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V.,
 Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
 Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
 Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
 Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
 McCann,R., Waterston,R. and Wilson,R.

REFERENCE Unpublished (1999)
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu

JOURNAL COMMENT This clone is available through: ResGen, Invitrogen Corp. 2130
 South Memorial Parkway Huntsville, AL 35801 For further information
 call: (800)-533-4363 or contact via email: ccu@resgen.com
 High quality sequence stop: 424.
 FEATURES Location/Qualifiers
 source 1. .565
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-c1084-1040"
 /tissue="Etiolated hypocotyls (Williams 82)"
 /lab_host="DH10B"

Query Match 6.2%; Score 76.6; DB 12; Length 565;
 Best Local Similarity 51.1%; Pred. No. 2.4e-05;
 Matches 179; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Qy 82 GTTACCCGGCTGACCTTACACTCCAGGCAATGCCCGGGAAACGCTCATGGGGCT 141
 Db 215 GTAACCAGGGTCTGTATACTGACAAGGATGTTAGGTGCAAGGTATGTAACAT 274

Qy 142 ATGAAAGGGCCCTTGAGGTTAGAACATCATCGGAACGACTCGGAAACATCATTGGCAAGT 201
 Db 275 ATGGAACCTGCTGGCTATCTGTGAGAAGGGATGCTGGTGTGG 334

Qy 202 GAAGGCAGTGTACCGAGCTTCCTGCGATCGGGTCAACACTCGATTCTGTCGA 261
 Db 335 GATGGCCCTTGAAACCTGAGGTTGCTGCAGTTGCAACAGGTTGCACATTGATGCTATACT 394

Qy 262 AACGGGGGATGTTGAGGGCTTGGGAGCTGGGGCTCTGAGGTGCTGGGGCT 321
 Db 395 TACTGGAAAATATGATGGAGTTGGGTGTTAGGTGCTATGAAGCCATCAGAGTC 454

Qy 322 ATGCTGGAGAACGGCTACGTGAATCGGATCCATTGAGTTCATCGGATCTGGGAGG 381
 Db 455 CTGAAAAGGTCTGGTTAAACCTAGAAGACCATTGGAAAGTCATTCACATCAGAA 514

Qy 382 GAAGGGGCCGCTTCAGGAGTGGCATGTTGGCCATTGGCAGG 431
 Db 515 GAACCAAACAGCTTGGAAATAGGTTGCTGGAAAGCCGCCTATGGCTGG 564

RESULT 15

LOCUS BX255493 669 bp mRNA linear EST 25-FEB-2003
 DEFINITION BX255493 Pinus pinaster differentiating xylem adult Pinus pinaster
 CDNA clone PP006D11, mRNA sequence.

ACCESSION BX255493
 VERSION BX255493.1 GI:28563495
 SOURCE EST.
 ORGANISM Pinus pinaster
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
 1 (bases 1 to 669)

REFERENCE Canton,F.R., Le Provost,G., Garcia,V., Barre,A., Frigerio,J.-M.,
 Paiva,J., Fevereiro,P., Avila,C., Mouret,J.-F., Brach,J., de
 Darvar,A., Canovas,F.M. and Plomion,C.

AUTHORS Dardar,
 TITLE Transcribedome analysis of wood formation in maritime pine
 JOURNAL Unpublished (2003)
 COMMENT Contact: Frigerio JM
 Genetique et Amelioration INRA
 route d'Arcachon 33612 Cestas CEDEX France
 Email: Frigerio@ pierrotton.inra.fr
 Location/Qualifiers

FEATURES

RESULT 12 Db 463 TCTGGAGGAGCAACGGACAGTCATTGGCGTCAACGCCATCGTGGGCAACGTGGTCA 522

REF ID AF029423 Qy 657 GCGGGTTGCCGTCAAAPGGCAGAAGGCACCGCCCAACCCCCTATGCCACCTGGGCCA 716

LOCUS DEFINITION Db 523 CACGGTGAACGGCTCAACGGGAGTCCAACATGCCAACGGCACAAACGCCATGGGTATCGTCG 582

Accession AF029423 Salmonella typhimurium LT2, Lambda DASH II Salmonella typhimurium genomic clone 1244-T3, genomic survey sequence.

Version AF029423 GI:2570953

Keywords GSS.

Organism Salmonella typhimurium

Salmonella typhimurium Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Salmonella.

Reference 1. (bases 1 to 925) Wong, R.M.-Y., Wong, K.K., Benson, N.R. and McClelland, M.

Authors Sample sequencing of a Salmonella typhimurium LT2 lambda library: comparison to the Escherichia coli K12 genome

Title FEMS Microbiol. Lett. 173 (2), 411-423 (1999)

Journal 99243757

Medline 10227170

Pubmed Contact: McClelland M

Molecular Biology
Sidney Kimmel Cancer Center
3099 Science Park Road, San Diego, CA 92121, USA

Email: mclelland@lifsci.sdsu.edu

Class: shotgun

Features Location/Qualifiers

source 1. .925 /organism="Salmonella typhimurium"
/mol_type="genomic DNA"
/strain="LT2"
/db_xref="taxon:602"
/clone_id="1244-T3"
/clone_lib="Salmonella typhimurium LT2, Lambda DASH II"
/note="Vector: Lambda DASH II; sequenced using Li-Cor sequencer"

Query Match 6.3%; Score 78.2; DB 28; Length 925;

Best Local Similarity 48.4%; Pred. No. 1.e-05;

Matches 281; Conservative 0; Mismatches 293; Indels 7; Gaps 2;

Db 177 ACTCGGAAACATCATGGCGGACTGTGAAAGGCACTGATCCGGAGCTTCCTGGATCGCGT 236

Db 50 ACGTGGAATTATAATGGCGCCCTTGGCACACAAATTCCGGAGCGGTATTAAAG 109

Db 237 CGGTTCACTTGCGATTCTGTCGAAACGGGGATGTTGATGGCACTGGCAGGGGTGTT 296

Db 110 CGGTCGATATTGATACTGGCTGTAACGGCAACGGCAATTGGACGGCAATTGGCGCT 169

Db 297 GTGGCCCTTGAGGCTGCCGGTATGCTGGAGAACGGCTACGTGAATCGGCATCCATT 356

Db 170 TGCGCTGGCTGACTGGCTCAAAGGACCTACGGCGCTGSACC-GT 228

Origin

Query Match 6.3%; Score 78.2; DB 28; Length 925;

Best Local Similarity 48.4%; Pred. No. 1.e-05;

Matches 281; Conservative 0; Mismatches 293; Indels 7; Gaps 2;

Db 417 CCGGGCATTGGAGGGTTGGCAGGGAACAGGGAACTGGACTCTTGGTTGATGAGGATGG 416

Db 229 CGAGGTAGTGTCAATGGCTGAGGAAGGGAGCCATTGGCTGGGAATCTGAAGAGGTACGCCATATCCAGGACGCCAAGG 288

Db 289 CAAGAACATTTGGCTGGCTGAGGCTATGGGATGGCTAAAGGACCTACGGCGCTGSACC-GT 348

Db 477 AGTGTCTGGTAGGAGGGCTACTGCCTTGGCTTAAGCCGGGAACCTGGCAGGGCTGC 536

Db 349 CACAGGCTTGTGAGGCTATGGGATGGCTAAAGGACCTACGGCGCTGSACC-GT 408

Db 537 AGCCCGCTCCGGGGACCTGGCTTATCGAAACTACATGGAAAGGACCGAT 596

Db 409 TGCGCGAACCTGGCTGGCTTGTGATGGCATATTGAGCAAGGCTGCGT 462

Db 597 CCTCGAGCAGGAAATAGAGATGGAGTGGTAACCTCCATCGTGGGTTGGCATT 656

ORIGIN

RESULT 13 BI498910 LOCUS BI498910.Y1 Gm-c1053 Glycine max cDNA clone GENOME SYSTEMS CLONE DEFINITION ID: Gm-c1053-3980 5' similar to TR:085664 O85664 N-CARBAMYL-L-AMINO ACID AMIDOHYDROLASE. ; mRNA sequence.

Accession BI498910

Version BI498910.1 GI:15338254

Keywords EST.

Source Glycine max (soybean)

Organism Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids II; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

1. (bases 1 to 514)

Reference Shoemaker, R., Keim, P., Vodkin, L., Erpelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

Title Public Soybean EST Project

Comment Unpublished (1999)

Contact: Shoemaker R/Public Soybean EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800) -533-4363 or contact via email: ccu@resgen.com

High quality sequence stop: 419.

Location/Qualifiers

1. .514 /organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="GENOME SYSTEMS CLONE ID: Gm-c1053-3980"

/tissue_type="whole seedling, 3 week old, greenhouse grown"

/lab_host="DH10B"

/clone_lib="Gm-c1053"

/note="Vector: Bluescript II SK+; Site_1: EcoRI; Site_2: XbaI; The Harosoy NIL was constructed and seed was provided by Dr. J. Specht, University of Nebraska (Shoemaker and Specht, 1995). The cDNA library was constructed from mRNA isolated from whole seedlings of 3 week old greenhouse grown plants. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XbaI restriction site and a 3' anchor. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XbaI digestion. The cDNA fragments were directionally cloned into the EcoRI-XbaI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in cooperation with Dr. Paul Keim's laboratory at Northern Arizona University."

ACCESSION	BZ566026	GI:27194755			
VERSION	BZ566026.1				
KEYWORDS	GSS.				
SOURCE	Pseudomonas aeruginosa				
ORGANISM	Pseudomonas aeruginosa				
REFERENCE	Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, B.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.				
AUTHORS	Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library				
TITLE	J. Bacteriol. (2002) In press				
JOURNAL	Contact : Chris K. Raymond				
COMMENT	Genome Center University of Washington Box 352145, Seattle, WA 98105-2145, USA Tel: 2062216954 Fax: 2066857244 Email: craymond@u.washington.edu				
FEATURES	shotgun				
	Location/Qualifiers				
	1. .885 /organism="Pseudomonas aeruginosa" /mol_type="genomic DNA" /strain="2-164" /db_xref="taxon:287" /clone="pac82-164 587" /clone_lib="pac82-164" /note="clinical isolate 2-164 Whole genomic shotgun library."				
ORIGIN					
	Query Match 6.8%; Score 83.8; DB 28; Length 885; Best Local Similarity 51.4%; Pred. No. 9.6e-07; Matches 242; Conservative 0; Mismatches 227; Indels 2; Gaps 2;				
Qy	686 ACGCCGGACAACCCCATGACCTGCCAGGATGCCGTGGTACCCGCCGCTCTCATGG 745				
Db	520 ACCGGGGCAAAACGAATTGCACCTCCCACGGGACCCCTGCTGGTGCCTGGCTTT 461				
Qy	746 TGAGGGAGGTCAACCGGTTCTGTCAACGAGATGCCGATGGCACAGTG-GCTACCGTTGGC 804				
Db	460 TCCAAGGGTTCAATCGGGAAAGCCCTCGGCCAATGCCGCAATGCCGATGGTACCCCTG 401				
Qy	805 CACCTCACAGTGGCCCCGGTGGAGGCAACCAAGTCCGGGGAGGGTGGACTTCACACTG 864				
Db	400 TGCCTGCACGCCATTGCCAACCGGACTGGATACCCGGCAAGTGAAGATGACCCCTG 341				
Qy	865 GACCTTGCGTTCTCGCATGAGGGAGTCGGCTCCGGTGTGATGCCACGGCATCTGGTCATG 924				
Db	340 GACTTCGCCATTGCCAACCGGAGGGCTGGACTGGATGCCGAGGGTCCGCCAGGTG 281				
Qy	925 GTCGGGAGGTGGCTCCAGGGCGGGTGGCTGGCGATGGCATTAATTCGAATTC 984				
Db	280 ATGCCCGCTACCTGGCGAACGATGGCTTGCAATACGAGCTGGTGCCTGGCTTTC 221				
Qy	985 AGCCCCGTTGCGAGTGGCTCCTAACATGGGGACGGCTGGAAAGGGCTCGGCCTTG 1044				
Db	220 CGGCCGCTRACTCGACCAAGGGATGGCTGGCGAACGGGGCTGGCTG 162				
Qy	1045 CAGTTCACACCCGGATATCAGGAGTGGGGGGCACGACTCCGATGTTCATGCCAG 1104				
Db	161 GGATGCCAGATGGACATCGTCAAGGGGATGATCTTCCTGCCAA 102				
Qy	1105 GTCACGGACGTCGGAATGGTTTCGTTCAAGCCGTGCTGGCCGGAGCCAC 1155				
Db	101 CTCGGTCCGGGGATGATCTTCGCTGCCAGGGCATCTGCCAC 51				
RESULT	CD443670				
LOCUS	572 bp	mRNA	linear	EST 03-JUN-2003	
Qy	697 ACCCCCCATGCCACCTGGCCAGGATGGCTGCTGGCCGGCT 7311				
Db	493 GTGCTGATGCCCTGCAAGAAATGCTGGACTGGCAGAGCT 534				

"library": "db_xref=taxon:287",
"clones": "pac81-60_1678",
"clone_id": "pac81-60",
"note": "clinical isolate 1-60 Whole genomic shotgun

HIGGIN

RESULT 9
 4422460
LOCUS BZ422460
DEFINITION id54a10.b1 WGS-SbicolorF (DH5a methyl filtered) Sorghum bicolor
ACCESSION BZ422460
VERSION BZ422460.1 GI:26370760
KEYWORDS GSS.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
DIFFERENCE 1 (bases 1 to 642)
AUTHORS Rabinowicz,P.D., O'Shaughnessy,A.L., Balija,V., Dedhia,N., Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L., Zutavern,T., Palmer,L., McCombie,W.R. and Martienssen,R.A.
TITLE Genomic shotgun sequences from Sorghum bicolor (methyl-filtered)
JOURNAL Unpublished (2002)
COMMENT Contact : W. Richard McCombie
 Lita Annenberg Hazen Genome Sequencing Center
 Cold Spring Harbor Laboratory
 PO Box 100, Cold Spring Harbor, NY 11724, USA
 Tel: 516 367 8884
 Fax: 516 367 8874

Email: mccoombie@cshl.org
Plate: id54 row: a column: 10
Seq primer: -21M13UnivFwd
Class: shotgun
High quality sequence stop: 642.

Éducation/Qualifiés

1. • 642
/organism="Sorghum bicolor"
/mol_type="genomic DNA"
/db_xref="taxon:4558"
/clone="id54a10"
/lab_host="DH5a"
/clone_lib="WGS-SbicolorF (DH5a methyl filtered)"
/note="Site 1: Xba I; Site 2: Xba I; The vector was digested with XbaI and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (.x/y reads in M13mp19.b/g reads in pUC19). The same ligation was transformed into DH5a."

Query	Match	7.3%	Score	90.6	DB	28	Length	642	
	Best Local Similarity	48.2%	Pred.	No.	3.4e-08				
Matches	286	Conservative	0	Mismatches	304	Indels	3	Gaps	1
Qy	593	CGATCCTCGAGGAGCAAATAAGAGATCGGAGTTGTAACCTCCATCGTGGCTTCGGCG	652						
Db	9	CGTTCTGCTCAATGAAGACCTGCCGTGGCATCGGCCATCTCGGCCAGGCC	68						
Qy	653	CATTGCCGGTGTGCCGTCAAAGGCCAGAACGGACCACGGCCACAAACCCCCATGGCACCTGC	712						
Db	69	GCTTCATGGCCGAGGTGCATGGCTTGCCATGGCCATGGCGAACCGTGC	128						
Qy	713	GCCAGGGATGGCTGGTACCCGGCTCATGGTGAGGGAGGTCAACCGGTTGGTCAAAC	772						
Db	129	GCGCGACGCCGCGCATGGCCGGAGATGGCCTACATGGCCATGGCGTGCAGC	187						
Qy	773	AGATGCCGATGGCACAGTGGTACCGGTTGGCCACAGTGGCCGGTGGAGGCA	832						
Db	188	-ATCAAGCCGGCTGGCCTGGCCACCATGGGCTTGGAGGTGGTGCAGGGGCCA	245						
Qy	833	ACCAAGTCCGGGGAGGTGGACTTCACACTGGACCTGGCGTTCTCCGCATGAGGAGTGTGC	892						
Db	246	ACGTGGTCCGGGACTGGCCAGCTGGCCATCGACATCCGGGAAGACGATGGCGAC	305						
Qy	893	TCCGGTGTGATCGACCGCATCTCGGTCAATGGTGGCGAGGGTCTCCCAGGGGGTG	952						
Db	306	GCCTGGCCAGTGGCCAGGTCAAGGGCGAGATCGAGGCATCGGGCATGGCCAGG	365						
Qy	953	TGGCTGCCGATGTGGATGAATTTCAAATCTCAGCCCCGGTGCAGCTGGCTCCTAACCATGG	101						
Db	366	TGCAGGTTGCATTGCCAGACCCATGAGGCCCATGAGGGCCAGGGCCTGGCCCTGGCGTCTGC	425						
Qy	1013	TGGACCGCGTTCGCGAAGCGGCCCTCGGCCCTGCAGTTCACACACCGGATATCAGGAGTG	107						
Db	426	AGGCCGGCTGGCAATTCCATTGGCCAGCCCCGGTGGCCACCTGCCATCTG	495						
Qy	1073	GGGGGGCCACCGACTCGATGTTCATGGCCAGGTACGGACGGTGGAAATGGTTTCGTTTC	113						
Db	486	GCGGGGCCACGGCCATGGCCCTGGGGCCTGACCGACGTGGCCATGCTGTCGTC	545						
Qy	1133	CAAGCCGTGCTGGCCGGAGCCACGTTCCCGAAGAATGGACCGATTTCGATGAC	1185						
Db	546	GCTGGGCAACGGGCATCAGCCATACCCGACGAGACCATGACCGAAAGGCC	598						

Best Local Similarity 48.8%; Pred. No. 8.5e-09; Matches 292; Conservative 0; Mismatches 294; Indels 12; Gaps 1;

QY 146 AAGCGGCCGCTTGGCTGAAAGACGCAACTCGGAAACATCATGGCCGACGTGAAG 205
DB 68 AGCGGCCGCTCACGATGGCTGACAAGATGGCACGATGTTGCCACCCGCCCG 127
QY 206 GCACTGATCCGGAGCTTCTGGATCGGGTACACTTGATTCTGTCCGAAACG 265
DB 128 GCGAGGACCCGGATGGCACTGGGCACTTGCAGGGTGGTGGCCCTTGAATGCCACCCAGGCCAACCG 187
QY 266 GCGGGATGTTGATGGCACTGGGCACTTGCAGGGTGGTGGCCCTTGAATGCCACCCAGGCCAACCG 325
DB 188 GCGGCAAGTTCGATGGCGTTCTGGCGTGGACTGGAAAGTGTGG 247
QY 326 TGGAAACGGCTACGTGAATGGCATCCATTGAGTTCATCGGATCGTGGAGGGAAAG 385
DB 248 ACGATCTAACATAAGCAGAAACATTGGTCTTACCTAACACTGGTCAA 307
QY 386 GGGCCCGCTTCAGCAGTGGCATGGGGGGCCATTGCAAGGGTTGGCTGCCCCGACA 445
DB 308 GCGCTCGCTTGGCCATGGCTGGCTTCCGGGATTCACGATCTCG 367
QY 446 GGGAAACTGGACTCTTGGTGTGAGGTGGAGATGGCTGGCTACTGGCT 505
DB 368 ATTACGCCCTATAGCCGACCGATAACGGCAAGACCTATGGGACGAACTGAAGGCCA 427
QY 506 TGGCTTGAAGCGGGGAACCTGCAAGGGCTCCGGGACTCTGGTGCCT 565
DB 428 TCGGCTGGCTGGTGAGGAAGGGTGGGGCC-----AAGATGCCGCT 475
QY 566 TTATCGAACTACATTGAAACAAGGACCGATCTCGAGCAGGAAATAGAGATCGGAG 625
DB 476 ATTTTCGAAATATCATATCGAACAGGGACCGATCTCGAGGGCAAGGAGATCGGCC 535
QY 626 TTGTAACCTCCATGTTGGCGATTGCGCAATTGGGGTGGCTGGAAAGTGGCTGAGCTGACGGCATCTCGGGT 685
DB 536 TCGTTACCCATGGTCAGGGCCTGTGGTGGCTGGAAAGTGAACGGCTGACGGCTGAGCTGACGGCT 595
QY 686 AGCGCGGACAAACCCCATGGCACCTGGCAAGGGATGGCTGGTACCCGGCTCTCAT 743
DB 596 ATACCGGCTCGACGCCATGGCCATGGCGTCAATGGCCATGGGGCTCGCCGCCAT 653

RESULT 8
BZ549118/c

BZ549118 1250 bp DNA linear GSS 17-DEC-2002
LOCUS pac81-60_1678.s1 pac81-60 Pseudomonas aeruginosa genomic clone
DEFINITION pac81-60_1678, genomic survey sequence.
ACCESSION BZ549118
VERSION BZ549118.1 GI:27152699
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonadaceae; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonas aeruginosa.
COMMENT (bases 1 to 1250)
REFERENCE Spencer, D.H., Raymond, C.K., Smith, B.E., Sims, E.E., Hastings, M.,
AUTHORS Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In Press
JOURNAL Contact: Chris K. Raymond
COMMENT University of Washington
Genome Center
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun
FEATURES Source
CLASS: Location/Qualifiers
1. .1522
/organism="Pseudomonas aeruginosa"
FEATURES Source
CLASS: Location/Qualifiers
1. .1250
/organism="genomic DNA"
/strain="1-60"

Db	389	CCAGGGATGGTGGGGGTTGGCGAGGGCCACGACTCGATGTTCATCGCCAGGTCA CGGACG-TCGGAA	448	Qy	503 CCTTCGGCTTGAAGCCGGGGAAC TGAGGCTGCAGCCCGTCCGGGGACCTGCGTG	562	
Qy	1062	TATCAGCAGTGGGGGGCCATTGATGTTCATCGCCAGGTCA CGGACG-TCGGAA	1120	Db	460 CCATCGGCT-----ACGGGGAGGGCGATCTGACATCGGATGGGGCG	413	
Db	449	CATCGTAGGGGCCACGACGCGATCTCCCGCGAACACTCGGTCCGGGAGA	508	Qy	563 CTTTTATCGAACTACACATTGAACAAAGGACCGATCCTCGAGGAGCAATAGAGATCG	622	
Qy	1121	TGGTTTCTGTTCCAAGCCGGTGTGGCCGGAGCCACGTTCCCGAAGAATGGACCGATTTCG	1180	Db	412 CCTATTCGAGGGCACATGGACAGGGCGATCCTCGAGGACGAGGAAAGGACATCG	353	
Db	509	TGATCTCGRGCCCCCTGCGAGAACGGCATCGCCACAACGAGATCGAGAACGCCAGGCCCG	568	Qy	623 GAGTTGTAACCTCCATCGTTGGGATTGCCGTCATAAGGCAGAAAGGCC	682	
Qy	1181	ATGACCT 1187		Db	352 GCGTGGTGTCTGGCGCTGGGAGAAATGGTTGACACTGTCCTGGCGTCGAGG	293	
Db	569	ACGACCT 575		Qy	683 ACCACGCCGGCACACCCCCATGACCTGCGCCAGGATGGCTGGTACCCGCCGATCTCA	742	
RESULT 5	BZ570771/c	BZ570771. msh2_1528.y2 genomic survey sequence.	1343 bp	DNA linear GSS 17-DEC-2002	Qy	292 CACACGCCGGCCAACGCCGATGACCTGCCGCTGGTCCCTGGCGTCGAGG	233
LOCUS	BZ570771. msh2_1528.y2	Pseudomonas aeruginosa genomic clone msh2_1528,			Db	743 TGGTGGGGAGGTCAACGGGTTGTCACAGGATGCCGATGGGCAAGTGGCTACCGGTTG	802
DEFINITION		genomic survey sequence.			Db	232 TGGTCGAGGGCAATCGGGCCATGCCGATGGGCACTGGCTTGCGGACGGTGC	173
ACCESSION	BZ570771	GI:27205832			Qy	803 GCCACCTCACAGTGGCCGGTGGAGGCAACCGGTCCGGGGAGGGTGGACATTCACAC	862
VERSION	BZ570771.1	GSS.			Db	172 GCTGCCCTGACGCCATCCGGTTCGGCCAACGTGATAACCGGGAAAGTGAAGATGACCC	113
SOURCE	Pseudomonas aeruginosa				Qy	863 TGGACCTGCGTTCTCGGCATGAGGAGTCGCT	893
ORGANISM	Pseudomonadaceae; Pseudomonas				Db	112 TGGACTTCGCCATCTGCAGGGGATCCACT	82
REFERENCE	Spencer, D.H., Raymond, C.K., Smith, B.E., Sims, E.B., Hastings, M., (bases 1 to 1343)			RESULT 6			
AUTHORS	Burns, J.L., Kaul, R. and Olsen, M.V.			CC130324			
TITLE	Whole-Genome-Sequence variation among multiple isolates of Psedomonas aeruginosa library			LOCUS	CC130324		
JOURNAL	J. Bacteriol. (2002) In press			DEFINITION	NDL:58C11.SP6 Notre Dame Aedes aegypti genomic clone		
COMMENT	Contact: Chris K. Raymond			ACCESSION	NDL:58C11, Genomic survey sequence.		
ORGANISM	Genome Center			VERSION	CC130324		
REFERENCE	Spencer, D.H., Raymond, C.K., Smith, B.E., Sims, E.B., Hastings, M., (bases 1 to 1343)			KEYWORDS	GI:29999379		
AUTHORS	Burns, J.L., Kaul, R. and Olsen, M.V.			SOURCE	GSS.		
TITLE	Whole-Genome-Sequence variation among multiple isolates of Psedomonas aeruginosa library			ORGANISM	Aedes aegypti		
JOURNAL	J. Bacteriol. (2002) In press			REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes; Stegomyia.		
COMMENT	Contact: Chris K. Raymond			AUTHORS	Loftus, B., Shetty, J., Knudson, D. and Severson, D.		
ORGANISM	Genome Center			TITLE	BAC end sequencing of Aedes aegypti		
REFERENCE	Spencer, D.H., Raymond, C.K., Smith, B.E., Sims, E.B., Hastings, M., (bases 1 to 1343)			JOURNAL	Unpublished (2003)		
AUTHORS	Burns, J.L., Kaul, R. and Olsen, M.V.			COMMENT	Other GSSs: NDL:58C11.T7		
TITLE	Whole-Genome-Sequence variation among multiple isolates of Psedomonas aeruginosa library			CONTACT	Contact: Brendan Loftus		
JOURNAL	J. Bacteriol. (2002) In press			DEPARTMENT	Department of Eukaryotic Genomics		
COMMENT	Contact: Chris K. Raymond			TIGR	9712 Medical Center Drive, Rockville, MD 20850, USA		
ORGANISM	Genome Center			TEL	Tel: 301-838-3543		
REFERENCE	Spencer, D.H., Raymond, C.K., Smith, B.E., Sims, E.B., Hastings, M., (bases 1 to 1343)			FAX	Fax: 301-838-0208		
AUTHORS	Burns, J.L., Kaul, R. and Olsen, M.V.			EMAIL	Email: enta@tigr.org		
TITLE	Whole-Genome-Sequence variation among multiple isolates of Psedomonas aeruginosa library			LIBRARY	Library was provided by David Severson		
JOURNAL	J. Bacteriol. (2002) In press			SEQ PRIMER	Seq primer: SP6		
COMMENT	Contact: Chris K. Raymond			CLASS	Class: BAC ends.		
FEATURES	Location/Qualifiers			FEATURES	Location/Qualifiers		
SOURCE	1. .1343 /organism="Pseudomonas aeruginosa"			SOURCE	1. .694 /organism="Aedes aegypti"		
	/mol_type="genomic DNA"				/mol_type="genomic DNA"		
	/strain="MSH"				/strain="Liverpool"		
	/db_xref="taxon:287"				/db_xref="taxon:7159"		
	/clone="msh2_1528"				/clone lib="Notre Dame Liverpool"		
	/clone_lib="msh"				/note="Vector: pECBAC1; Site 1: Hind III; The library was prepared from whole body tissue of newly hatched L1 larvae by David Severson at the University of Notre Dame and Hongbin Zhang"		
ORIGIN	/note="Environmental isolate. Whole genomic shotgun library."						
Query Match	7.7% ; Score 95.4 ; DB 28 ; Length 1343 ; Best Local Similarity 49.4% ; Pred. No. 4.8e-09 ; Matches 312 ; Conservative 0 ; Mismatches 306 ; Indels 13 ; Gaps 2 ;						
Db	263 ACGGGGATGGTGTGATGGCACTGAGGCTGGCCCTGAGGTGCGCCGGTGA 322						
Qy	699 ACCGGTCAAGTTGACGGCTGGCTTCGGGTGATGCCGCTCGAGGTGA 640						
Db	323 TGCTGGAGAACGGCTACGTGAATCGGCATCCATTGAGTTCATCGGCTGGAGG 382						
Qy	639 CTCACGAAACCTCGGGTGGAAACCCAAGGGCGCTGAAGGTGGTGTGGACCC 580						
Db	383 AAGGGGGCCGCTTACGGAGTCATGGTGGGGCCATGCAAGGGTGGTGC 442						
Qy	579 AAGGCTGGCCTTGCGCCCTGCAATGGGGTGGTGGTGGTCAACCC 520						
Db	443 ACAGGGAACTGGACTCTTGGTTGATGAGGATGGAGTGGTGTGGCTACTG 502						
Qy	519 TGAGGAGAC-CCTGGCCAAGGGGATGCCGAGGGTGTAGGGAGGGCTGGACAGGGT 461						
Db	Query Match 7.6% ; Score 93.6 ; DB 28 ; Length 694 ;						

SEQUENCE	CA757386	Qy	685 CACGCCGGACAACCCCATGCGCACCTGGCCAGGATGGCTGGTACCCGGCTTCATG 744
VERSION	CA757386.1	Db	541 CATGCTGGAGCAGTGGCTCATGCTGAAGAAATGATGCTGGACTGGCTGAGTTG 600
KEYWORDS	EST.		
SOURCE	Oryza Sativa	Qy	745 G 745
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae; Oryza.	Db	601 G 601
REFERENCE	Bohnert, H.J., Borchert, C., Brazille, S., Brooks, J., Baton, M., Ferrea, H., Kawasaki, S., McCollough, A., Michalowski, C.B., Palacio, C., Scarpa, G., Wheeler, M. and Zepeda, G.R.	RESULT 4 BZ567409	1226 bp DNA linear GSS 17-DEC-2002
AUTHORS	Functional Genomics of Plant Stress Tolerance	LOCUS	BZ567409
TITLE	Unpublished (2000)	DEFINITION	pacs2-164_7024_x1 pacs2-164 Pseudomonas aeruginosa genomic clone
JOURNAL	COMMENT	ACCESSION	pacs2-164_7024_genomic survey sequence.
COMMENT	Contact: Mark Fredrickson	VERSION	BZ567409
FEATURES	Department of Plant Biology	KEYWORDS	GI:27198622
source	University of Illinois	ORGANISM	GSS.
	1201 W. Gregory Dr., Urbana, IL 61801, USA	SOURCE	Pseudomonas aeruginosa
	Tel: 2172655473	ORGANISM	Pseudomonadaceae; Pseudomonas
	Email: bohnertlab@life.uiuc.edu.	REFERENCE	Pseudomonas aeruginosa
		AUTHORS	Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M., Burns, J.L., Kaul, R. and Olsen, M.V.
		TITLE	Whole-Genome-Sequence variation among multiple isolates of Pseudomonas aeruginosa library
		JOURNAL	J. Bacteriol. (2002) In press
		COMMENT	Contact: Chris K. Raymond
			Genome Center University of Washington Box 352145, Seattle, WA 98105-2145, USA
			Tel: 2062216954
			Fax: 2066857244
			Email: craymond@u.washington.edu
			Class: shotgun.
			Location/Qualifiers
			1. .1226
			/organism="Pseudomonas aeruginosa"
			/mol type="genomic DNA"
			/strain="2-164"
			/db_xref="taxon:287"
			/clone="pacs2-164_7024"
			/clone_lib="pac52-164"
			/note="Clinical isolate 2-164 Whole genomic shotgun library."
ORIGIN			
	Query Match 7.7%; Score 95.8; DB 14; Length 667;		
	Best Local Similarity 49.9%; Pred. No. 2.9e-09;		
	Matches 300; Conservative 0; Mismatches 292; Indels 9; Gaps 2;		
		FEATURES	
		source	
	154 GCCTTGAGCGTTCTGTGAAGAGCAGCACTCGGAAACATCATCGGCCGACGTGAT 213		
	1 GGCTGTCCATTCTGTGAAGAGCAGCTGGTGGCAACATTTTGGGGCTGGAGGGTTCTGAG 60		
	214 CGGGAGCTTCCCTGGATCGGGTCAACTTCGATTCTGTCGAAACCGGGGGATG 273		
	61 GCAGGACTAGGAGCGGGTGTGAACTGGTCTCATGTTGACGCAATCCATTCTCAGGCAA 120		
	274 TTGATGGCACTGCAGGGTGGTGTGGCCGGCTTGAGGGCTGCCGGGTGATGCTGGAAAC 333		
	121 TATGATGGGTGGTGGCTGGTGGCTGGTGGCTGGTGGCAATTAGAATGCTGAAAGGTCT 180		
	334 GGCTACGTGAATCGGATCCATTGAGTTCATCGGATCGTGGAGGAGGGCCCGC 393		
	181 GGTTTCCAGGCCAAAAGATCTTGGAGGTATTATGTTACATCGGAGGCCATAAGA 240		
	394 TTCAAGGAGTGGCATGTTGGGGCCATTGAGGTTCATCGGATCGTGGTGGTGGTGG 450		
	241 TTTGGAAATTAGCTGCTTGGAAAGCCGCTTAATGGCAGGGAGTGAAGAACTGGCTCGATCC 300		
	451 CTGGACTCTTTGAAAGCAGGGGAAGTGGAGGTGGTCCGGTGTGGCTGGTGGCTGG 504		
	301 CTCAAAGAACAGTTGACAATTCAGAATGTTGTCATTTTGTGATGCTGGTGGC 360		
	505 TTCTGGGCTTGAAAGCAGGGGAAGTGGAGGTGGCTGGCTGGGGGACCTGGCTGGTGGCT 564		
	361 TACAAGATGCATCTGAGGAGCTACATGTTGTCATTTGTGTTTGTGATGCTGGTGGC 420		
	565 TTTATGAACTACATTGAAACAAGGACCGATCCTCGAGGAGCAAATAGAGATCGGA 624		
	421 TTGTTGAAATTGCACTTGGCCATTGGAAAGAGTCAAAATTGGC 480		
	625 GTTGTAACTCCATCGTGGCTTGCGCATGGGGTTGCCGCTCAAGGAGAAGGAC 684		
	481 GTTGTCACTGCTATTGCTGCTCCTGCAAGTATTAAAGTGGAGTTGAAGGAATGGGGC 540		
	Qy	1002 TCCTACCATGGTGGAAAGGGCCGTTGCGAAGTCACACACGGGGTACAGTTCA 1061	

1. .935
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="1-60"
/db_xref="taxon:287"
/clone="pac81-60"
/clone_lib="pac81-60"
/note="clinical isolate 1-60 whole genomic shotgun library."

ORIGIN

Query Match 10.2%; Score 125.8; DB 28; Length 935;
 Best Local Similarity 52.1%; Pred. No. 2.3e-15;
 Matches 280; Conservatve 0; Mismatches 257; Indels 0; Gaps 0;

QY 685 CACGCCGGCACACCCCCATGCACTGGCCAGGATGGCCTGGTACCCGCCTCTCATG 744
 DB 772 CAAGCCGGGCCAACGCCGATGCCCTGGCAATGATGCCCTGGTCGTTGCCGGCTTG 713

QY 745 GTGAGGGAGGTCAACGGTTCAACGAGATGCCGATGGCACAGTGGTACCGTTGCG 804
 DB 712 CTCGAGGGGTCAATGCCGCTGCCCTCGCCATCAGCCGATGCTTGGCACGGTGGC 653

QY 805 CACCTAACAGTGGCCCGGGGAGGTGGGACTTCACACTG 864
 DB 652 TGCCTGACGCCATATCCGGTTGCCGAAACGTGATAACCCGGGAAGTGAAGATGACCCCTG 593

ORIGIN

Query Match 10.2%; Score 125.8; DB 28; Length 935;
 Best Local Similarity 52.1%; Pred. No. 2.3e-15;
 Matches 280; Conservatve 0; Mismatches 257; Indels 0; Gaps 0;

QY 685 GACCTGGTTCCTCGGCATGAGGGAGTCGCTGATGCCGATCTCGGTCTATG 924
 DB 592 GACTTCGGCCATCTGCAACGGAGGCCCTGGACTCGATGATGCCGAGGTCCAGGTRG 533

QY 925 GTCGGGAGGTGGCTCCAGGGGGTGTGGCTGCCGATGGGATGAATTTCAAATCTC 984
 DB 532 ATCGCCGCTACCTGGAGAAGCATGGCTGCAATAAGGAGCTGGCTGGTGGCGACGCCGATTTC 473

QY 985 AGCCCCGGTGGCAGCTCCTTACCATGGGGCTTCGGGAAGGGGCTCGGCCCTTG 1044
 DB 472 CGGCCGCTGTACTCGACCAGGGATGCTGGCGGGTGGCAGGGCAGGGCTG 413

QY 1045 CAGTTAACACACCGGGATATCAGCAGTGGGGCCACGACTCGATGTTCATGCCCAAG 1104
 DB 412 GGATGCCGAGATGGACATCTGTCAGGCCAGCGGATCTTCCTCGCCGAA 353

QY 1105 GTCACGGACGTGGATGGTTTCAAGCCGGTGGCTGCCGAGCCACGGTTCGGCAA 1164
 DB 352 CTGGTCCGGGGATGATCTTGTGCCCTGGAGAACGGCATCAGCCACAACGAGATC 293

QY 1165 GAATGGACCGGATTCTGATGACCCCTTGCAAGGAACCTGAGGTTGTCCTCCGGTAATG 1221
 DB 292 GAGAACGCCAGGCCAGCGACCTGGCCGGCTGGGGTGTGTTGGGGCCATG 236

QY 1221 GAGAACGCCAGGCCAGCGACCTGGCCGGCTGGGGTGTGTTGGGGCCATG 236

Cold Spring Harbor Laboratory

PO Box 100, Cold Spring Harbor, NY 11724, USA

Tel: 516 367 8884

Fax: 516 367 8874

Email: mccombie@cshl.org

Plate: hr41 row: h column: 1.2

Seq primer: -21M13UnivRev

Class: shotgun

High quality sequence stop: 614.

FEATURES

source

1. .614
/organism="Sorghum bicolor"
/mol_type="genomic DNA"

/xref="taxon:4558"

/clone="hr41h12"

/lab_host="JM107 or DH5a"

/clone_lib="WGS-SbicolorF (JM107 adapted methyl filtered)"

/note="Site 1: Xba I; Site 2: Xba I; The vector was

digested with XbaI and one nucleotide was added by fill in
in the recessive 3' end. The genomic DNA was nebulized,
end repaired, adaptor ligated and size fractionated using
sephadex. The resulting fragments were between 0.8 and 3
kb and were cloned into the vector (.x/y reads in M13mp19,
.b/g reads in pUC19). The same ligation was transformed in
either JM107 or DH5a."

Location/Qualifiers

1. .614
/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/xref="taxon:4558"

/clone="hr41h12"

/lab_host="JM107 or DH5a"

/clone_lib="WGS-SbicolorF (JM107 adapted methyl filtered)"

/note="Site 1: Xba I; Site 2: Xba I; The vector was

digested with XbaI and one nucleotide was added by fill in
in the recessive 3' end. The genomic DNA was nebulized,
end repaired, adaptor ligated and size fractionated using
sephadex. The resulting fragments were between 0.8 and 3
kb and were cloned into the vector (.x/y reads in M13mp19,
.b/g reads in pUC19). The same ligation was transformed in
either JM107 or DH5a."

Location/Qualifiers

1. .614
/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/xref="taxon:4558"

/clone="hr41h12"

/lab_host="JM107 or DH5a"

/clone_lib="WGS-SbicolorF (JM107 adapted methyl filtered)"

/note="Site 1: Xba I; Site 2: Xba I; The vector was

digested with XbaI and one nucleotide was added by fill in
in the recessive 3' end. The genomic DNA was nebulized,
end repaired, adaptor ligated and size fractionated using
sephadex. The resulting fragments were between 0.8 and 3
kb and were cloned into the vector (.x/y reads in M13mp19,
.b/g reads in pUC19). The same ligation was transformed in
either JM107 or DH5a."

Location/Qualifiers

1. .614
/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/xref="taxon:4558"

/clone="hr41h12"

/lab_host="JM107 or DH5a"

/clone_lib="WGS-SbicolorF (JM107 adapted methyl filtered)"

/note="Site 1: Xba I; Site 2: Xba I; The vector was

digested with XbaI and one nucleotide was added by fill in
in the recessive 3' end. The genomic DNA was nebulized,
end repaired, adaptor ligated and size fractionated using
sephadex. The resulting fragments were between 0.8 and 3
kb and were cloned into the vector (.x/y reads in M13mp19,
.b/g reads in pUC19). The same ligation was transformed in
either JM107 or DH5a."

Location/Qualifiers

1. .614
/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/xref="taxon:4558"

/clone="hr41h12"

/lab_host="JM107 or DH5a"

/clone_lib="WGS-SbicolorF (JM107 adapted methyl filtered)"

Location/Qualifiers

1. .614
/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/xref="taxon:4558"

/clone="hr41h12"

/lab_host="JM107 or DH5a"

/clone_lib="WGS-SbicolorF (JM107 adapted methyl filtered)"

Location/Qualifiers

1. .614
/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/xref="taxon:4558"

/clone="hr41h12"

/lab_host="JM107 or DH5a"

/clone_lib="WGS-SbicolorF (JM107 adapted methyl filtered)"

Location/Qualifiers

1. .614
/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/xref="taxon:4558"

/clone="hr41h12"

/lab_host="JM107 or DH5a"

/clone_lib="WGS-SbicolorF (JM107 adapted methyl filtered)"

Location/Qualifiers

1. .614
/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/xref="taxon:4558"

/clone="hr41h12"

/lab_host="JM107 or DH5a"

/clone_lib="WGS-SbicolorF (JM107 adapted methyl filtered)"

Location/Qualifiers

1. .614
/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/xref="taxon:4558"

/clone="hr41h12"

/lab_host="JM107 or DH5a"

/clone_lib="WGS-SbicolorF (JM107 adapted methyl filtered)"

Location/Qualifiers

1. .614
/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/xref="taxon:4558"

/clone="hr41h12"

/lab_host="JM107 or DH5a"

/clone_lib="WGS-SbicolorF (JM107 adapted methyl filtered)"

Location/Qualifiers

1. .614
/organism="Sorghum bicolor"

/mol_type="genomic DNA"

/xref="taxon:4558"

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 13:15:23 ; Search time 3685 Seconds
 (without alignments)
 10040.500 Million cell updates/sec

Title: US-10-045-063-1
 Perfect score: 1239
 Sequence: 1 gtgaccctgcagaaaaggcga.....tgaaggccattgaccggtaa 1239

Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 550265578

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing First 45 summaries

Database : EST:
 1: em_estba:
 2: em_estbu:
 3: em_estin:
 4: em_estmu:
 5: em_estov:
 6: em_estp:
 7: em_estro:
 8: em_htc:
 9: gb_est1:
 10: gb_est2:
 11: gb_htc:
 12: gb_est3:
 13: gb_est4:
 14: gb_est5:
 15: em_estfun:
 16: em_estom:
 17: em_gss_hum:
 18: em_gss_inv:
 19: em_gss_pln:
 20: em_gss_vrt:
 21: em_gss_fun:
 22: em_gss_mam:
 23: em_gss_mus:
 24: em_gss_pro:
 25: em_gss_rod:
 26: em_gss_phg:
 27: em_gss_vrl:
 28: gb_gss1:
 29: gb_gss2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
C 1	125.8	10.2	935 BZ549119	BZ549119 pac81-60
C 2	96.6	7.8	614 BZ349536	BZ349536 hr41h12.9
C 3	95.8	7.7	667 CA757386	CA757386 BR0500120
C 4	95.8	7.7	1226 BZ567409	BZ567409 pac82-164

FEATURES

RESULT 1
 BZ549119/c
 LOCUS pac81-60_1678.82 pac81-60 Pseudomonas aeruginosa genomic clone
 DEFINITION pac81-60_1678, genomic survey sequence.
 ACCESSION BZ549119
 VERSION 1
 KEYWORDS GSS.
 SOURCE Pseudomonas aeruginosa
 ORGANISM Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 Pseudomonadaceae; Pseudomonas.
 REFERENCE 1 (bases 1 to 935)
 AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
 Burns, J.L., Kaul, R. and Olsen, M.V.
 TITLE Whole-Genome-Sequence variation among multiple isolates of
 Pseudomonas aeruginosa library
 J. Bacteriol. (2002) In press
 JOURNAL Contact: Chris K. Raymond
 COMMENT Genome Center
 University of Washington
 Box 352145, Seattle, WA 98105-2145, USA
 Tel: 2062216954
 Fax: 2066857244
 Email: raymond@u.washington.edu
 Class: shotgun.
 Location/Qualifiers

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
C 1	125.8	10.2	935 BZ549119	BZ549119 pac81-60
C 2	96.6	7.8	614 BZ349536	BZ349536 hr41h12.9
C 3	95.8	7.7	667 CA757386	CA757386 BR0500120
C 4	95.8	7.7	1226 BZ567409	BZ567409 pac82-164

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 4, 2004, 12:20:58 ; Search time 5119 Seconds
 (without alignments)
 10490.725 Million cell updates/sec

Title: US-10-045-063-1

Perfect score: 1239

Sequence: 1 gtgaccctgtcagaaaggcga.....tgaaggcacccggtaa 1239

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing First 45 summaries

Database : GenEmbl:
 1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_p1:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_p1:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rod:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_by:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

Score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1239	100.0	1239	6	AR198375	Sequence AR198375 Sequence AX015596 Sequence BD194451 Arthrobac Sequence AX103678 Sequence AX370705 Sequence AF146701 Arthrobac Sequence AX370708 Sequence AX397881 Sequence BD176614 DNA encod BD180889 DNA encod BD180889 DNA encod BD181025 5-Subsit BD176615 DNA encod BD180890 DNA encod BD181026 5-Subsit AL591975 Listeria Sequence AX641665 Sequence AP005957 Bradyrhiz AE012124 Xanthomon C 18 222.6 18.0 11322 1 AE012124 Xanthomon C 19 222.6 18.0 11322 1 AE012124 Xanthomon C 20 220.6 17.8 249050 1 AL591975 Listeria Sequence AX417038 Sequence AX417041 Sequence AE004481 Pseudomon C 21 220.6 17.8 349980 6 AX417038 Sequence AX417041 Sequence AE004481 Pseudomon C 22 220.6 17.8 349980 6 AX417041 Sequence AE004481 Pseudomon C 23 214.8 17.3 11276 1 AE004481 Pseudomon C 24 210.8 17.0 1380 6 AR385713 Sequence AE011655 Xanthomon C 25 207.6 16.8 13275 1 AE011655 Xanthomon C 26 205.6 16.6 348997 1 BX640427 Bordetell C 27 205.6 16.6 349008 1 BX640444 Bordetell C 28 194.4 15.7 10813 1 AE005928 Caulobact AE016789 Pseudomon AE003012 Mesorhizo BX640445 Bordetell BX640446 Bordetell AE016776 Pseudomon AP001509 Bacillus AF425838 Geobacill E03775 DNA sequenc S67784 N-carbamyl- AP002997 Mesorhizo BX640439 Bordetell BX640425 Bordetell Y08752 B.stearothe Y08751 B.stearothe BSORF4 AP004604 Oceanobac BX640417 Bordetell BX569695 Synechoco

ALIGNMENTS

RESULT 1	AR198375	LOCUS	AR198375 Sequence 1 from patent US 635284.8.
DEFINITION	Sequence 1	ACCESSION	AR198375 Sequence AR198375.1 GI:20248224
KEYWORDS		KEYWORDS	
SOURCE		SOURCE	Unknown.
ORGANISM		ORGANISM	Unclassified.
			1 (base 1 to 1239)
REFERENCE		AUTHORS	Altenbuchner,J., Matthes,R., Pietzsch,M., Syldatk,C., Wiese,A. and Wilms,B.
TITLE			Recombinant L-N-carbamoylase from Arthrobacter aurescens and method of producing L-amino acids therewith

JOURNAL	Patent: US 6352848-A 1 05-MAR-2002;	Db	901 CTGATCGACCGCATCTCGGTATGGTCAGGCCAGGGCTCCAGCCGGTGGCTGCC 960	
FEATURES	Location/Qualifiers	Qy	961 GATGTGGATGAATTTCATCTAGCCCCGGTGCAGTCCTACCATGGTGGACGCC 1020	
source	1. 1239 /organism="unknown" /mol_type="unassigned DNA"	Db	961 GATGTGGATGAATTTCATCTAGCCCCGGTGCAGTCCTACCATGGTGGACGCC 1020	
ORIGIN		Qy	1021 GTTCGGAAAGGGCCTGGCCCTTGCAGTTACACACCGGATATCAGCAGTGGGGGGC 1080	
	Query Match 100.0%; Score 1239; DB 6; Length 1239; Best Local Similarity 100.0%; Pred. No. 8.9e-208; Matches 1239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Db	1021 GTTCGGAAAGGGCCTGGCCCTTGCAGTTACACACCGGATATCAGCAGTGGGGGGC 1080	
Qy	1 GTGACCTCTGCAGAAAGGCCAAGGGGCCATTGAGAAAGAGATCCGGGAGCTCTCCGG 60	Qy	1081 CACGACTCGATGTTCATCGGCCAGGTACCGGACGTGGAAATGGTTTCAGTCCAAAGCCGT 1140	
Db	1 GTGACCTCTGCAGAAAGGCCAAGGGGCCATTGAGAAAGAGATCCGGGAGCTCTCCGG 60	Db	1081 CACGACTCGATGTTCATCGGCCAGGTACCGGACGTGGAAATGGTTTCAGTCCAAAGCCGT 1140	
Qy	61 TTCTCGGCAGAAAGGCCCGGTTACCCGGTGTACCTACACTCAGGCAATGCCGCCGG 120	Qy	1141 GCTGGCCGAGGCCACGTTCCCGAAGAATGGGATTCAGCTTGCAAAAGGAACCT 1200	
Db	61 TTCTCGGCAGAAAGGCCCGGTTACCCGGTGTACCTACACTCAGGCAATGCCGCCGG 120	Db	1141 GCTGGCCGAGGCCACGTTCCCGAAGAATGGGATTCAGCTTGCAAAAGGAACCT 1200	
Qy	121 CGGGAAACGCTCATTCGGGCTATGAAAGGGCCGCTTGAGCTTCTGTGAAGACGCACTCAGGCAATGCCGCCGG 120	Qy	1201 GAGGTGTCCTCGGGTAATGAAGGCACTTGACCGGTAA 1239	
Db	121 CGGGAAACGCTCATTCGGGCTATGAAAGGGCCGCTTGAGCTTCTGTGAAGACGCACTCAGGCAATGCCGCCGG 120	Db	1201 GAGGTGTCCTCGGGTAATGAAGGCACTTGACCGGTAA 1239	
Qy	181 GAAAACATCATGGCCGACGTGAAGGCCACTGATCCGGAGCTCTGGATCGGGTGGT 240	RESULT 2		
Db	181 GAAAACATCATGGCCGACGTGAAGGCCACTGATCCGGAGCTCTGGATCGGGTGGT 240	AX015596	1239 bp	DNA
Qy	241 TCACACTTCGATTCTGTTCGAAACGGGGATGTTGATGGCACTGCAGGGGTGGTGC 300	LOCUS	Sequence 4 from Patent WO9951722.	PAT 07-SEP-2000
Db	241 TCACACTTCGATTCTGTTCGAAACGGGGATGTTGATGGCACTGCAGGGGTGGTGC 300	DEFINITION	AX015596	
Qy	301 GCCCTTGGGCTGCCCGGGTATGCTGGAAACGGCTAACGGCTACATCCATTGAG 360	VERSION	AX015596.1 GI:10041431	
Db	301 GCCCTTGGGCTGCCCGGGTATGCTGGAAACGGGACTGTGCACTGGCATCCATTGAG 360	KEYWORDS	.	
Qy	361 TTCAATCGCGATCGTGGAGGAGGAAGGGCCGCTTCAAGCAGTGGGGGGGG 420	SOURCE	Arthrobacter aurescens	
Db	361 TTCAATCGCGATCGTGGAGGAGGAAGGGCCGCTTCAAGCAGTGGGGGGGG 420	ORGANISM	Bacteria; Arthrobacter aurescens	
Qy	421 GCCATTGGCAGGGTTGGTCGCCGACGGGAACGGGAACTGGACTCTTGGTTGATGAGGGATGGAGTG 480	REFERENCE	Micrococcaceae; Arthrobacter.	
Db	421 GCCATTGGCAGGGTTGGTCGCCGACGGGAACGGGAACTGGACTCTTGGTTGATGAGGGATGGAGTG 480	AUTHORS	Syldatk,C., Mattes,R., Wilms,B., Pietzsch,M., Wiese,A. and Altenbuchner,J.	
Qy	481 TCCGTTAGGCAGGGGCTACTGCCTTGCTTATCGAACTACACATGGAAAGGGCTGCC 540	TITLE	Recombinant 1-n-carbamoylase derived from arthrobacter aurescens, and a method for producing 1-amino acids by using the same	
Db	481 TCCGTTAGGCAGGGGCTACTGCCTTGCTTGAAGGGGGAAACTGCAGGGCTGCC 540	JOURNAL	Patent: WO 9951722-A 4 14-OCT-1999; DEGUSSA (DE); UNIV STUTTGART (DE); ROCHE DIAGNOSTICS GMBH (DE)	
Qy	541 CGCTCCGGGGGAACTGCGTGGCTTATCGAACTACACATGGAAAGGGCTGCC 600	FEATURES	Location/Qualifiers	
Db	541 CGCTCCGGGGGAACTGCGTGGCTTATCGAACTACACATGGAAAGGGCTGCC 600	Source	1. 1239	
Qy	601 GAGCAGGAGCAAATAAGATGGAGCTGGAGTTGTAACCTCCATCGTGGGGCATGGGG 660	transl_table	/organism="Arthrobacter aurescens"	
Db	601 GAGCAGGAGCAAATAAGATGGAGCTGGAGTTGTAACCTCCATCGTGGGGCATGGGG 660	protein_id	/mol_type="unassigned DNA"	
Qy	661 GRTGCCGTCAAAGGAGAAAGGACCAAGGCCAACGGGCAAAACCCCCATGCACTTGGCCAGGAT 720	db_xref	/db_xref="GI:10041432"	
Db	661 GRTGCCGTCAAAGGAGAAAGGACCAAGGCCAACGGGCAAAACCCCCATGCACTTGGCCAGGAT 720	translation	/db_xref="REMTRMBL:CAC07595"	
Qy	721 GGCCTGGTACCCGGCGCTCTATGGTGAAGGGTCAACGGGTCAACGAGATGCC 780	db_xref	/translation="MTLOKAQAAARIKEIRELSRFSAEAGPGVTRLTYPPEHAARETL EAAMKAAALSVREDALGNIIGRREGTDPELPAIVGSHFDVSVRNGGMFDGTVGVCAL SYRQAATAFGLKPGEHQAAARSAAIDLRAFIELHIEQGPILQEQQIEIGVVTSLVGVR ALRVAVKGRSIDHAGTTPMHLRQDAVPAALMVRREVNRFVNIEIADGTVATVGHLTVAPEG GNQVPGEVDFTLDRSPHEESLRVLIDRISVMVGEVASQAGVAADVDEFNLSPVQLA PTMVDVREASALQFTHRDISSGAGHDMSMFIAQVTDMVFVPSRAGRSHVPEEWTD FDDLRKGTEVLRVMKALDR"	
Db	721 GGCCTGGTACCCGGCGCTCTATGGTGAAGGGTCAACGGGTCAACGAGATGCC 780	ORIGIN		
Qy	781 GATGGCACAGTGGCTACCGTGGCCACCTCACAGTGGCCGGTGGAGGAAACAGGTC 840	Query	Score 1239; DB 6; Length 1239;	
Db	781 GATGGCACAGTGGCTACCGTGGCCACCTCACAGTGGCCGGTGGAGGATGCC 840	Match	Best Local Similarity 100.0%; Pred. No. 8.9e-208;	
Qy	841 CCGGGGGAGGTCAAGCTGGCATGGAGACTCGCTCCGGTCTCCGGT 900	Mismatches	0; Mismatches 0; Indels 0; Gaps 0;	
Db	841 CCGGGGGAGGTCAAGCTGGCATGGAGACTCGCTCCGGTCTCCGGT 900	Conservative		
Qy	901 CTGATCGACCGCATCTCGGTATGGTCAGGCCGGTGGCTGCC 960	1	GTGACCCCTGGCAGAAAGGCCAACGGCTACCTGGCATGGAGCTCCGGTCTCCGG 60	
		1	GTGACCCCTGGCAGAAAGGCCAACGGCTACCTGGCATGGAGCTCCGGTCTCCGG 60	
		61	TTCCTCGGCGAGAAAGGCCAACGGCTACCTGGCATGGAGCTCCGGTCTCCGG 120	

PLIITRRSVERDYAETVEQTVIALLEGIEEDTRFDVLDSSIVPLPLAVVTKLIGFDAED
 " complement (478. · 1884)
 /note="hypothetical protein"
 /codon_start=1
 /evidence=not experimental
 /transl_table=11
 /product="transport protein homolog"
 /protein_id="AAG02126.1"
 /db_xref="GI : 9931309"
 /translation="MKEESAAAEVYOGASFIPDEORSGRPRDLFFMFSLSANVTOLYVVI
 GATIMVFGLSMWQAVVIVILVANSILMFLMGWASIGAASGVPTSIIITRALFGMRGNKVF
 GAGVVVVLSTLVITVNLVIGTQGGLVVAERLGMAPDVTTAIVCGIIVLASTGLAIYG
 YMMVVKTSQCVLTIVIIAAAFVLLGFQKVDSLASFQPVPTENPVIGFIVALTLCIAMP
 LSWGSAADEPARHIPSTTRRKPIVLAVGGLMMIPLVGLGEVGLVIAATTATDMSEPVTLGIM
 NLVPDWFGTIVLWVVIVIGSIMNTMITYGGSLLAVOAGGPVSRRVVAUTVGLGLTAIGT
 AYVLFDFADSFNLQLGLTCLGSVAMPLAIVLVDVRKGRYDGVALSDPSQGGPYW
 HRGCFRLISGGVALLSUNGASYQSPIAALLGGIDMSIVLGGTVAAVLFWAL
 SRQEERATSNKLTERTPQ"
 complement (1893. · 2114)
 /note="hypothetical protein"
 /codon_start=1
 /evidence=not experimental
 /transl_table=11
 /product="ferredoxin homolog"
 /protein_id="AAG02127.1"
 /db_xref="GI : 9931310"
 /translation="MRFVFDQIKCTGHVCEELAPDAFLIEDGIAIRVLLIAEVHGATL
 AEVQAACQCPHQAITLAEPDPSSSAENA"
 2358. · 3794
 /gene="hyuP"
 2358. · 3794
 /gene="hyuP"
 /codon_start=1
 /transl_table=11
 /product="permease homolog HyuP"
 /protein_id="AAG02128.1"
 /db_xref="GI : 9931311"
 /translation="MTTTEDDQGGRSLNNESNAPTRYAQRSVGPFSILAAIWFMAMAIQVA
 IFIAAGQLTASFQWWQVIVAAAGCTIAAVVVFPTQSAAIRWGINFTVAARMFPGIRG
 SLIPIVVKALLSLFWFGFQTWLGAALADEITRILTGFNSNLTLWVLLIFGAQVQVTTFYG
 ISFIRWMNIFAAAPVLLAMGVFMVYMLNGAVTGLGEVMSMGGETPGMPFSTAVMIFVG
 GWIAVVVSITHDIVEKECKVSPSESTREGRKSDAKYAAAQWGMVPASAVIFGFIGGAASMV
 LYGEWNPNPVEAITKWKVMPQFADTLLNSFLNVLSPAYTLCSTPPRVF
 SFKTGIVVSGAIGLMMMPQFADTLLNSFLNVLASALGPLAGIMIDYFLVRRGRVLSQ
 DLYRTKGIVTYWHGFNWVIALAVYVVALAFLTPHLMFITGLVAAVVLHYPMARWVK
 SFPMFSEAASENEEYLEQEDGELATHQ"
 3921. · 4631
 /gene="hyuA"
 /codon_start=1
 /transl_table=11
 /product="hydantoin racemase HyuA"
 /protein_id="AAG02129.1"
 /db_xref="GI : 9931312"
 /translation="MRLIVVINPNSSSSALTESVADAQQVVTATGT1ISAINPSRGPAVI
 EGSFDEALATEFHLLIEVERAERENPPDAYVIACFGDPGLDAVKELTDPRVVGVAEAAI
 HMSSFVAATFSIVSILPRVKHLIELVRQAGATNRLASIKLPNLGMMAFHEDHEAHE
 TIKQAKEAVQEDGAEASIVLGCAGMVGFAQRQLSDELGPVIDPVEAACRVAESLVALG
 YOTSKANSYQKPTEKQYL"
 4651. · 6027
 /gene="hyuH"
 4651. · 6027
 /gene="hyuH"
 /codon_start=1
 /transl_table=11
 /product="L-hydantoinase HyuH"
 /protein_id="AAG02130.1"
 /db_xref="GI : 9931313"
 /translation="MFDDVIVKNCRMVSSDGTIEADILVKDGKVAIASADTRDVEASRT
 IDAGGCKFVMMPGVVDEHVH1IDMDLKNRYGRFELDSESAVGGITTEIMPITFPPTT
 LDASFLEKKKQAGQRLKVDFALEYGGGPVNLPFIRKMDAGAVGFKSMMMAASVPGMFDAA

RESULT 6
 AF146701 LOCUS AF146701 DNA linear BCT 04-OCT-2000
 DEFINITION Arthrobacter aurescens cytochrome P-450 homolog, partial cds;
 and transport protein homolog, ferredoxin homolog, permease homolog
 HyuP (hyuP), hydantoin racemase HyuA (hyuA), L-hydantoinase HyuH
 (hyuH), L-N-carbamoylase HyuC (hyuC), and lacI repressor homolog
 genes, complete cds.
 AF146701 AF071221
 AF146701.1 GI:9931307
 KEYWORDS Arthrobacter aurescens
 ORGANISM Bacteria; Actinobacteria; Actinomycetales;
 Micrococcineae; Micrococcaceae; Arthrobacter.
 SOURCE Wilms,B., Wiese,A., Syldatk,C., Matthes,R., Altenbuchner,J. and
 Pietzsch,M.
 TITLE Cloning, nucleotide sequence and expression of a new
 L-N-carbamoylase gene from Arthrobacter aurescens DSM 3747 in E.
 coli
 JOURNAL J. Biotechnol. 68 (2-3), 101-113 (1999)
 MEDLINE 99210756
 PUBMED 10194852
 PUBLICATION 2 (bases 1 to 8715)
 AUTHORS Wiese,A., Pietzsch,M., Syldatk,C., Matthes,R. and Altenbuchner,J.
 TITLE Hydantoin racemase from Arthrobacter aurescens DSM 3747;
 heterologous expression, purification and characterization
 J. Biotechnol. 80 (3), 217-230 (2000)
 JOURNAL 20403588
 MEDLINE 10949312
 PUBMED 3 (bases 6003 to 7322)
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1998) Institute of Industrial Genetics,
 University of Stuttgart, Allmandring 31, Stuttgart 70569, Germany
 REFERENCE 4 (bases 1 to 8715)
 AUTHORS Wiese,A.H. and Altenbuchner,J.
 TITLE Direct Submission
 JOURNAL Submitted (29-APR-1999) Institute of Industrial Genetics,
 University of

Sequence Data							
Accession	Organism	Description	Length	Source	Comments		
VSDGELFEI FQEIAACGSVIVVHAENETI IQALQKQIKAAGGKDMAAYEASQPVFQEN EAIQRALLJOKEAGCRLILHVSNPDGVELIHQAQSEGDVHCESGPQYLNITTDDAE RIGPMKVAPPVRSAMNVRLEQLENGVITLGSDHGGHPVDEDKEPGWKDVMKAGNG ALCGLETSILPMMILNTGVNKGRSLSERLVEWMCEKPALKFLGIYPKQGTLOQVGSDDADLIL DLDIDTKYDASQFRSLHKYSPFDGMPVTGAPVLTMRGTVVAEQGEVLVEQGFQFVTRHYEASK"							
6024 . 7262 /gene="hyuc" 6024 . 7262 /gene="hyuc" /codon_start=1 /transl_table=11 /product="L-N-carbamoylase Hyuc" /protein_id="AAG02131_1" /db_xref="GI: 9931314" /translation="MTLQKAQQAERIEKIRELSRSAEGPGVTRLYTBEHAARETL IAAMKAAALSVERDALGNIGRREGTDPELPAIVGSHFDSVRNGMFDGTAGVVCAL EAARVMLENGYVNRHPPFEFIAVEEGARFSSGMIGRAJAGLVADREILSLVDDEGV SVRQAAATAFGLKPGELQAAARSAAIDLRAFIELHIEQGPILEEQIEGVVTISVGVR LRVAVKGRSDHAGTTPEMLRQDALVPAALMREVNRFVNEIADGTVATVGHLTVAPEG GNQVPGEVDFTLDLRSPHEESLRLVIDRISVMVGEVASQAGVAADVDEFNLSPVQLA PTMVDAVREASALQFTTHRSDGAGHDMSMFIAQVTDVGMVFVPSRAGRSHVPEEWTD FDDLRLKGTEVVLRVMKALDR" complement (7416 . 8450) /note="hypothetical protein" /codon_start=1 /evidence-not_experimental /transl_table=11 /product="lacI repressor homolog" /protein_id="AAG02132_1" /db_xref="GI: 9931315" /translation="MHMTEQLKRATISDVAEHPGVSRRAVSKVLRNAYGVSPAMRTKV EESMTLLNYRPLALARGMRGNSYTLGFLVDSLNTFFSVLAEIGIRDVAEELGYQVLIG QARQGLDAQQKLIEAMVDRRMGDLVLIAPFGPNDELEKLGRTIIPVWVLGRHGPAKYFD TVASDDIAGSALTVDHLSVLGLHERISMLANTQDPDSEPMGPQTVRHRGYDQAMEKHGLA KYIDDHVPGRWGTESGLEAGRSLAARPLAQRPSALHGGADVAALGAYTALFEAGIHIPD DMSIVGYDNIPATGLPPIGLSTVDQSGVQMGROQAELLSSRIEGRTESEHVLVSPRLI VRKTSAPPRR"							
CDS							
Query Match 99.9%; Score 1237.4; DB 1; Length 8715; Best Local Similarity 99.9%; Pred. No. 1.3e-207; Matches 1238; Conservative 0; Mismatches 1; Indels 0; Gaps 0;							
Qy 1 GTGACCTGCAAGAACGGCAAGGGGGCATTGAGAAAGAGATCCGGGAGCTCTCCGG 60 Db 6024 GTGACCTGCAAGAACGGCAAGGGAGCGCATTCGGAAAGAGATCCGGCTCTCCGG 6083	Qy 1081 CACGACTCGATGTTCATGCCCAAGGTACGGACCTGGAAATGGTTCCAAAGCGT 1140 Db 7104 CACGACTCGATGTTCATGCCCAAGGTACGGACCTGGAAATGGTTCCAAAGCGT 7163	Qy 1141 GCTGGCCGGAGCCACGTTCCCGAAGAAATGGACCTTGCAAAAGGAAC 1200 Db 7164 GCTGGCCGGAGCCACGTTCCCGAAGAAATGGACCTTGCAAAAGGAAC 7223	Qy 1201 GAGGTGTCCTCCGGTAATGAAGGCACCTGACCGGTAA 1239 Db 7224 GAGGTGTCCTCCGGTAATGAAGGCACCTGACCGGTAA 7262	RESULT 7 AX370708 LOCUS AX370708 DEFINITION Sequence 6 from Patent WO2010424. ACCESSION AX370708 VERSION AX370708.1 KEYWORDS GI:19168865 SOURCE Arthrobacter aurescens ORGANISM Arthrobacter aurescens REFERENCE 1 AUTHORS Krimmer, H.P., May, O., Klement, I., Drauz, K. and Reichert, D. TITLE Process for the preparation of allysine acetal JOURNAL Patent: WO 2010424-A 6 07-FEB-2002; Degerussa AG (DE)	FEATURES Location/Qualifiers		
ORIGIN							
Query Match 99.9%; Score 1237.4; DB 1; Length 8715; Best Local Similarity 99.9%; Pred. No. 1.3e-207; Matches 1238; Conservative 0; Mismatches 1; Indels 0; Gaps 0;							
Qy 1 CGGAAACATCATCGGCCGACGTGAAGGCACGTGATCCGGAGCTTCCCTGCGATCGGGT 240 Db 6084 TTCTGGCAGAAGGCCCGGGTACCGGCTGACCTACCTCCAGGGCATGCGCC 120 Db 6144 CGGAAACGCTCATCGGCCGACGTGAGGACTACCTCCAGGGCATGCGCC 6143	Qy 121 CGGAAACATCATCGGCCGACGTGAAGGCACGTGATCCGGAGCTTCCCTGCGATCGGGT 180 Db 6144 CGGAAACGCTCATCGGCCGACGTGAGGACTACCTCCAGGGCATGCGCC 6203	Qy 181 GGAAACATCATCGGCCGACGTGAAGGCACGTGATCCGGAGCTTCCCTGCGATCGGGT 240 Db 6204 GGAAACATCATCGGCCGACGTGAAGGCACGTGATCCGGAGCTTCCCTGCGATCGGGT 1263	Qy 241 TCACACTTCGATTCTGTCGAAACGGGGGATGTTGATGGCACTGCAGGCGTGGTGC 300 Db 6264 TCACACTTCGATTCTGTCGAAACGGGGGATGTTGATGGCACTGCAGGCGTGGTGC 6323	Qy 301 GCCCTTGAGGCTGCCGGGTGATGGAGAACGGCTACGTGAATGGCATCCATTGAG 360 Db 6324 GCCCTTGAGGCTGCCGGGTGATGGAGAACGGCTACGTGAATGGCATCCATTGAG 6383	Qy 361 TTCACTCGGATCGTGGAGGAAAGGGCCGCTTCAGCAGTGGCATGTTGGCGCCGG 420 Db 6384 TTCACTCGGATCGTGGAGGAAAGGGCCGCTTCAGCAGTGGCATGTTGGCGCCGG 6443	Qy 421 GCCATTGCAAGGGTTGGCTCGCCGACAGGGAACTGGACTCTTGGTTGATGGAGTGGAGTS 420 Db 421 QY	QY

	SOURCE	ORGANISM	COMMENT
Db	Microbacterium liquefaciens	Microbacterium liquefaciens	
Qy	Bacteria; Actinobacteridae; Actinomycetales;	Micrococcineae; Microbacteriaceae; Microbacterium.	
Db	Micrococcineae; Microbacteriaceae; Microbacterium.		
Qy	1 (bases 1 to 1239)		
Db	Takenaka, Y., Suzuki, S., Onishi, N. and Yokozeki, K.		
Qy	DNA encoding hydantoinase, DNA encoding N-carbamyl-L-amino acid		
Db	hydrolase, recombinant DNA, transformed cells, process for		
Qy	producing protein and process for producing optically active amino acid		
Db	Patent: JP 2002330785-A 2 19-NOV-2002;	AJINOMOTO CO INC	
Qy	PN JP 2002330785-A/2	OS Microbacterium liquefaciens	
Db	PD 19-NOV-2002		
Qy	PF 27-SEP-2001 JP 2001298619	PI YASUHIRO TAKENAKA, SHUNICHI SUZUKI, NORIMASA ONISHI, KENZO PI	
Db	PC C12N15/09, C12N1/15, C12N1/19, C12N5/10, C12N9/78, C12P13/22,	YOKOZEKI	
Qy	PC C12P41/00 // (C12N1/21, C12R1:01), (C12N9/78, C12R1:01), (C12P41/00, C12R1:01),	PC C12N15/00, C12NS/00	
Db	PC C12N15/00, C12NS/00	CC Chase	
Qy	FT 1239	FH Key	
Db	Location/Qualifiers	Location/Qualifiers	
Qy	1. organism="Microbacterium liquefaciens"	1. (1239).	
Db	/mol type="genomic DNA"		
Qy	/db_xref="taxon:33918"		
	FEATURES	Source	
Qy	ORIGIN		
Db	Query Match 59.7%; Score 739.8; DB 6; Length 1239;		
Qy	Best Local Similarity 74.8%; Pred. No. 4e-120;		
Db	Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;		
Qy	1 GTGACCTTGAGAAAGGCCAACGGCGCCATTGAGAAAGAGATCAGGGAGCTCTCCGG	Qy	1
Db	1 GTGACGCTGCAGCAGGGCGGGCGATCGCATCGGACTCTCCGG	Db	1
Qy	661 TTCTCGGAGAAGGCCCGGTGTACCCGGTGTACCTCCAGCTTCAACTCCAGGATCGCC	Qy	61
Db	661 ATCACGGTGGAGATGGGATCGTGTGACGGGATCGTGTGACGGGAGCTGGGAGGAC	Db	61
Qy	721 GGCCTGTAACCCCCGGCTCTCATGGTGAGGGTCAACCGGGTCTGTCAACGGAGATGCC	Qy	121
Db	721 GGCCTGTTGGGGGCTCATGGTGAGAGATCAAATGGTTGTGCAACGGAGATGCC	Db	121
Qy	781 GATGGCACAGTGGGTACCGTGGGCCACCTCACGGTGGGGGAAACCGAGTC	Qy	181
Db	781 GACGGCACGGTGGGACCGTGGGCCACCTCACGGTGGGGCTAACCGAGTT	Db	181
Qy	841 CCGGGGGAGGTGGACTTCACACTGGACCTCGGTCTCCGATCGCGGTG 900	Db	121
Db	841 CCCGGGGGGCTCGAGTTCAAGCTCGATCCGGCTGATCCGGCTC 900	Db	121
Qy	901 CTGATGACCCGATCTCGGTCAAGCTCGATCCGGCTCAGGGAGTCGGTGGCTGCC	Qy	1080
Db	901 CTGGTCAACAGGGATCGAGGGCATGGTGGCAGAAGTCGCCACCTACCGTGGACGCC	Db	1080
Qy	961 GATGTGGATGAATTTCACACTTCAGCCGGATGGTGGCTCTACCCATGGTGGACGCC	Qy	1020
Db	961 GCGGTGAACGGGGTCTCGGGCTCAGCCCTGTGGTGGTGGATGCC	Db	1020
Qy	1021 GTCAGCTGGCTCGGGCTTGGAGTCAGTCAACACCCGGATATCAGGAGTGGGGGGC	Qy	1140
Db	1021 GTCAGCTGGCTCGGGCTCAAGCTGGCTCAGCCATCGGACATCGAGGGGGAGGG	Db	1140
Qy	1081 CACGACTCGATGTTCATGCCAGCTGGCTCAAGCTGGCTGGAGTCGGTGGCTGCC	Qy	1200
Db	1081 CACGACTCGATGTTCATGCCAGATGCCAGCTGGCTGGAGTCGGTGGCTGCC	Db	1200
Qy	1141 GCTGGCGGAGCCACGTTCCCGAAGAATGGACCGATTTCGATGACCTTCGAAAGGA	Qy	1200
Db	1141 GCCGGGGAAAGCCATGGTCCGGAGGAATGGGGGAGCTGGTGAACGGGAGC	Db	1200
Qy	1201 GAGGTGTCCTCCGGTAATGAAAGGCACCTGACCGGTAA 123 9	Qy	1239
Db	1201 GATGTGGTCCCTCACGTCGTGACGGCGCTTGAACGGGTGA 123 9	Db	1239
RESULT 10	1239 bp DNA linear PAT 15-MAY-2003	Qy	421
BD180889	DNA encoding hydantoinase, DNA encoding N-carbamyl-L-amino acid	Db	421
DEFINITION	hydrolase, recombinant DNA, transformed cells, process for		
	producing protein and process for producing optically active amino acid.		
ACCESSION	BD180889	Qy	481
VERSION	BD180889.1 GR:30791807	Db	481
KEYWORDS	JP 2002330785-A/2.		

QY	541 CGCTCCGGGGGACCTGGTGTCTTGAACTACACATTGAAACAAGGACCGATCCTC 600	PI SHUNICHI SUZUKI, NORIMASA ONISHI, KENZO YOKOZEKI PC
Db	541 CGTACGAGGGATGACCTTCGGCCTCATCGAGTTGCACATCGAGCGGCCGATCCRC 600	C12N15/09, C12N1/21, C12N9/90, C12P13/02, C12P17/10, PC
QY	601 GAGCAGGAAANTAGAGATGGAGTTGTAACCTCATCGTTGGCTTCGGCATTTGGG 660	C12N15/00
Db	601 GAGCAGGAAAGGGATGGGGATCGGCTGTCGGATCGTCCGGTCTTCGGG 660	CC Chase
QY	661 GTTGGCGTCAAAGGCGAACACGGGACCAACCCCCCATGCACCTTGGCCAGGGAT 720	FEATURES
Db	661 ATCACGGTGGAGGGGACACGGGACCAACCCCCATGCACCTTGGCCAGGGAC 720	source
QY	721 GCGCTGGTACCCGGGCTCATGGTGAAGGGAGGTCAACGGTCAACGAGATGCC 780	ORIGIN
Db	721 GCGCTGGTGGCGGGCTCATGGTGGAGAGATCAATGGTTGTCAAAGAGATGGG 780	Query Match
QY	781 GATGGCACAGTGGTACCGTGGCCACCTCACAGTGGCCCAGGGAAACCAAGGGTC 840	Best Local Similarity
Db	781 GACGGCACGGTGGACCCCTCACGGTGAACCTGATGGAGTCGCTCCGGCT 840	Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;
QY	841 CCGGGGAGGGTGGACTTCACACTGGACCTGGGTTCTCGCAATGGAGTCGCTCCGGTG 900	QY
Db	841 CCCGGGGGCTGAGTTCAAGCTCGATCTGGCATGGCCACAGGGAGTCGATCCGGCT 900	Db
QY	901 CTGATCGACCGCATCTCGGTATGGTGGCTCCAGGGCGGTGCTGGCTGCC 960	QY
Db	901 CTGGTCGACAGGGATCGAGGGGATGGCAAGATGGCTGGCTGCCAGGGCC 960	Db
QY	961 GATGTCGATGAATTTCAAATCTCAGCCGGCATCTACCATGGTGGACGCC 1020	QY
Db	961 GCGGTGAACGGGTCTTCGGCTCAAGCCCTGTCGCTGTGGATTCGC 1020	Db
QY	1021 GTTCGGGAAGGGGCTCGGCTTGCACTGGTCAACACCCGGATATCAGCAGTGGGGC 1080	QY
Db	1021 GTGGGGGACGGGTCTTCGGCTCGGAACTCGGCTCACCCATCGGACATCACAGGGGGAGGG 1080	Db
QY	1081 CACGACTCGATGTTCATGCCCAAGGTCAAGGACGGTCACTGGTCCAAGCCGT 1140	QY
Db	1081 CACGACTCGATGTTCATGCCCAAGATCACCGGAGTCAGTGGCTGCCAGGCC 1140	Db
QY	1141 GCTGGCCGGAGGCCACGGTCCCGAAGGAAATGGACCGATTTCGATGACCTTCGCAAAGGAACCT 1200	QY
Db	1141 GCCGGGGAAAGCCATGTGGCCGGAGGAATGGTCCGAAAGGGGAGC 1200	Db
QY	1201 GAGGTTGTCCTCGGGTAATGAAGGCCACCTGGACCGGTAA 1239	QY
Db	1201 GATGTGGTCCTCACGTGTRGACGGCGCTTGAACGGATCTGGTGAACCCCTGGACACGGCGTG 1239	Db
RESULT 11		QY
BD181025	BD181025 1239 bp DNA linear PAT 15-MAY-2003	Db
DEFINITION	5-Substituted hydantoin racemase, DNA encoding the same, recombinant DNA, transformed cells and process for producing optically active amino acid.	QY
ACCESSION	BD181025 GI:30791943	Db
VERSION	JP 2002330784-A/3.	QY
KEYWORDS	Microbacterium liquefaciens	Db
SOURCE	Microbacterium liquefaciens	QY
ORGANISM	Bacteria: Actinobacteridae; Actinomycetales; Microbacterium, Micrococcineae; Microbacterium liquefaciens	Db
REFERENCE	1 (bases 1 to 1239)	QY
AUTHORS	Suzuki, S., Onishi, N. and Yokozeki, K.	Db
TITLE	5-Substituted hydantoin racemase, DNA encoding the same, recombinant DNA, transformed cells and process for producing optically active amino acid.	QY
JOURNAL	Patent: JP 2002330784-A 3 19-NOV-2002; AJINOMOTO CO INC	Db
COMMENT	OS Microbacterium liquefaciens	QY
	PN JP 2002330784-A/3	Db
	PD 19-NOV-2002	QY
	PF 13-SEP-2001 JP 2001278739	Db

hydrolase, recombinant DNA, transformed cells, process for producing protein and process for producing optically active amino acid.

ACCESSION VERSION	BD176615 BD176615.1 GI:29122325	Db	2465 GTCATGGGATCGTGAAGAGGGACCCGGCTTACGAGGGCATGGCTGC 2524
KEYWORDS	Microbacterium liquefaciens	Qy	421 GCCATTGAGGGTTGGCTGCCAACGGAAACTGGACTTTGGTGTAGGGATGGAGTG 4 80
SOURCE	Microbacterium liquefaciens	Db	2525 GCGATGGGGGCTGTCGCCGATCTGGACACCTGGACAGGGCGTGC 2584
ORGANISM	Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales; Micrococcineae; Microbacterium; Microbacterium.	Qy	481 TCCGTTAGGCAGGGCTACTGCCTTCGGCTTAAGCAGGGCAACTGAGGCTGCAGGC 540
REFERENCE	1 (bases 1 to 3343)	Db	2585 ACGTTGGGAGGGGCCACGGCTTGGGCTGGAAACCGGGTGAAGCTGGGC 2644
AUTHORS	Takenaka, Y., Suzuki, S., Onishi, N. and Yokozeki, K.	Qy	541 CGCTCCGGGGGACTCTGGGTGCTTTATGAACTACATGAAACAAGGACCGATCCTC 600
TITLE	DNA encoding hydantoinase, DNA encoding N-carbamyl-L-amino acid hydrolyase, recombinant DNA, transformed cells, process for producing protein and process for producing optically active amino acid	Db	2645 CGTACGAGGGATGACCTTCGGGCTTCATCGAGTTGACATCGAGGCACTGGATCCTC 2704
JOURNAL	Patent: WO 02072841-A 3 19-SEP-2002; AJINOMOTO CO INC, YASUHIRO TAKENAKA, SHUNICHI SUZUKI, NORIMASA ONISHI, KENZO YOKOZEKI	Qy	601 GAGCAGGZAGCAAATAGAGATGGAGTTGTAACCTCCATCGTTGGCTTCGGCATGGCG 660
COMMENT	OS Microbacterium liquefaciens PN WO 02072841-A/3 PD 19-SEP-2002 PF 08-MAR-2002 WO 2002JP002173 PR 08-MAR-2001 JP 01P 065814, 27-SEP-2001 JP 01P 298619 PI YASUHIRO TAKENAKA, SHUNICHI SUZUKI, NORIMASA ONISHI, KENZO PI YOKOZEKI PC C12N15/55, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N9/14, C12N9/86, C12P13/04, C12P13/22, C12P41/00, C12N15/52 CC DNA encoding hydantoinase, DNA encoding N-carbamyl-L-amino acid hydrolase,	Db	2705 GAGCAGGZAGCAAAGGTGAGATGGCTGACGGGGATCGCTGGCTCCGG 2764
FEATURES	CC recombinant DNA, transformed cells, process for producing CC protein and process for producing optically active amino acid PH Key Location/Qualifiers FT source 1. .3343 /organism='Microbacterium liquefaciens' FT Location/Qualifiers 1. .3343 /organism="Microbacterium liquefaciens" /mol_type="genomic DNA" /db_xref="taxon:33918"	Db	2765 ATCACGGTTGGAGGGCAGGGCACCACGGGGACCAACCCCCATGCACCTGGCCAGGAT 720
FEATURES source	Query Match 5.9.7; Score 739.8; DB 6; Length 3343; Best Local Similarity 74.8%; Pred. No. 3.5e-120; Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;	Db	721 GCGCTGGTACCCGGCTCATGGTGAAGGTCAACGGTTCTGTCACGAGATCGGC 780
ORIGIN	1 GTGACCCCTGAGAAAGGCCAACGGGGCATTTGAAAGAGATCGGGAGCTCTCCGG 60 2105 GTGACCGCTGAGCGGGCCGATCGCATCGAGGAGCTCTGGACTCTCCTCCGC 2164	Qy	2825 GCGCTGGTGGCCGGCTCATGGTGAAGGTCAACGGATCGACCTGGCTCAACGAGATCGGC 2884
	61 TTCTCGGAGAAAGCCCCGGTTACCCGGCTGACTACACTCCAGAGCATGCCGCC 120 2165 TTCTCGGTGAAAGGGCCGGCTGACAGTCTCAGTACACTCCGGAGCACGCCGCC 2224	Db	781 GATGGCACAGTGGACTTCAACCTGGACCTCGCTCATGGAGACTCGCTGGCTGC 840
	121 CGGGAAACGCTCATTTGGGGCATGAAAGCGGGCGCTTGTGAAAGACGCACTC 180 2225 CGAGAGGTGATCGTGGCGCATGGCGGAGGGCTGAGGGTCCACGAGGACGGCTC 2284	Qy	2945 CCCGGGGCTCGAGTTCACTGGGATCTGGGATCGCTGGCTCAACGGCT 3004
	181 GGAAACATCATCGGCCGACGGCACTGATCGGAGCTTCTGGATCGGGCTGGT 240 2285 GCGAACATCATCGTCTGGCTGGGGAGCCCGCTCTGGCGATGCCCTTGGC 2344	Db	901 CTGATCGAACGGCATCTCGGTCAGGGCTCCAGGGCTGGCTGC 960
	61 TTCTCGGAGAAACGGGGATGTTGATGGCACTGAGGCTGCTGGATGGTTTCGTTCAAGGGCT 1020 2345 TGGCACTCGACTGGCTCCGCAACGGGGATGTTGACGGGACGGCTGGTGC 3004	Qy	3005 CTGGTCGACAGGATCGAGGGAAGTGGCAGAACGGCTGGGAGTCGAGGCGC 3064
	1201 GAGGTTGTCCTCGGGTAATGAAGGCACCTGACCGGTAA 1239 3305 GATGGTCCCTCACGTGTCAGGGCTGACGGCTGACCGGGTA 3343	Db	961 GATGGGATGAAATTTCAAATCTCAGGCCGCTTACCATGGTGGACGCC 1020
	1021 GTTCGGGAAGGGCTCGGCCCTGGCATTGAGTTAACACGGGATATAGGAGTGGGGGG 1080 3125 GTGGCGGAGCGGGGCTCGGAACACTCGGCTTCAACCCATCGGACATCGAGGGGG 3184	Qy	3065 GCGGTGAACCGGGTTCTTCGGGCTCATGGCTGTCGGTCTGGATCGC 3124
	1081 CACGACTCGATGTTCATGCCAGGTACGGACGTGGATGGTTTCGTTCAAGGGCT 1140 3185 CACGACTCGATGTTCATGCCAGGTACGGACGTGGATGGTTTCGTTCAAGGGCT 1140	Db	1081 CACGACTCGATGTTCATGCCAGGTACGGACGTGGATGGTTTCGTTCAAGGGCT 1140
	3185 CACGACTCGATGTTCATGCCAGGTACGGACGTGGATGGTTTCGTTCAAGGGCT 1140 3245 GCGGGGGGAGGCCATGTGGCTGGGAATGGTCCGGAATTGCGAAGGGGAGC 3304	Qy	3185 CACGACTCGATGTTCATGCCAGGTACGGACGTGGATGGTTTCGTTCAAGGGCT 1140
	3245 GCGGGGGGAGGCCATGTGGCTGGGAATGGTCCGGAATTGCGAAGGGGAGC 3304 3305 GATGGTCCCTCACGTGTCAGGGCTGACGGCTGACCGGGTA 3343	Db	3245 GCGGGGGGAGGCCATGTGGCTGGGAATGGTCCGGAATTGCGAAGGGGAGC 3304
	RESULT 14 BD180890	Qy	3305 GATGGTCCCTCACGTGTCAGGGCTGACGGCTGACCGGGTA 3343
	LOCUS DEFINITION	Qy	3343 bp DNA linear
	DEFINITION	Qy	PAT 15-MAY-2003
	KEYWORDS	Qy	DNA encoding N-carbamyl-L-amino acid
	SOURCE	Db	DNA encoding hydantoinase, DNA recombinant DNA, transformed cells, process for producing protein and process for producing optically active amino acid.
	ORGANISM	Qy	BD180890
	ACCESSION VERSION	Qy	BD180890.1 GI:30791808
	KEYWORDS	Db	JP 2002330785-A/3
	SOURCE	Db	Microbacterium liquefaciens
	ORGANISM	Qy	Microbacterium liquefaciens
	ACCESSION VERSION	Qy	Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales;
	KEYWORDS	Db	Actinomycetales;

REFERENCE	Micrococcineae; Microbacteriaceae; Microbacterium. 1 (bases 1 to 3343)	2585	ACGGTGGCGAGGGGGCCACAGGCCCTTGGGCTGGAAACCGGGTGAGGTGGCTGGGACGGGGCC 2644
AUTHORS	Takenaka, Y.; Suzuki, S.; Onishi, N. and Yokozeki, K.	541	CGCTCCGGGGGACCTGGTATCGAACCTACACATTGAACAAGGACCGATCCTC 600
TITLE	DNA encoding hydantoinase, DNA encoding N-carbamyl-L-amino acid hydrolase, recombinant DNA, transformed cells, process for producing protein and process for producing optically active amino acid	2645	CGTAGCAGGGATGACCTTCGGCCTCATCGAGTTGCACATCGAGCAAGGGGATCCTC 2704
JOURNAL	Patent: JP 2002330785-A 3 19-NOV-2002;	601	GAGCAGGAGCAAATAAGAGATCGGAGTTGTAACCTCCATGTTGGCATTGGCG 660
COMMENT	AJINOMOTO CO INC OS Microbacterium liquefaciens	2705	GAGCAGGAGAAAGTGGAGATGGCGTGTGACGGGATCTGGCTTCGG 2764
PN	JP 2002330785-A/3	661	GTGCCGTCAAAGGAGAAAGGACCAACCCCCTATGCAACCTGGCCAGGAT 720
PD	19-NOV-2002	2765	ATCACGGTGGAGGGCAGGACCGCCGACGACGGGACTGCACCTGGCAGGAC 2824
PF	27-SEP-2001 JP 2001298619	721	GCGCTGGTACCCGGCTCATGGTGGCTCAACGGATCGTCAGCGAATGCC 780
PI	YASUHIRO TAKENAKA, SHUNICHI SUZUKI, NORIMASA ONISHI, KENZO PI YOKOZEKI	2825	GCGCTGGTGGCGGCTCATGGTGGAGAGATCAATGGGTCACCGAGATCGCG 2884
PC	C12N15/09, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N9/78, C12P13/22,	781	GATGGCACAGTGGCTTACCGTTCACACTGGACTCTGGCCACTCACAGTGGCCCCGGTGGAGGGCAACCAGGTC 840
PC	C12P41/00// (C12N1/21, C12R1:01), (C12N9/78, C12R1:01), (C12P41/00, C12R1:01),	2885	GACGGCACGGTGGCGACCTCACGGTGAACCCCTGGTGGAGAGATCAACCGAGATCGCG 2944
PC	C12N15/00, C12N5/00	841	CGGGGGAGGGTGGGACTTCACACTGGACTCTGGCTCCGGTGGAGGTG 900
CC	DNA encoding hydantoinase, DNA encoding N-carbamyl-L-amino acid hydrolase, recombinant DNA, transformed cells, process for producing CC protein and process for producing optically active amino acid FH Key	2945	CCGGGGGGCGTGCAGGTTCAAGGCTCGATCTGGGATCGCTCGATCCGGCTC 3004
CC	Location/Qualifiers	901	CTGATCGACCCGATCTCGGTATCGGCTCCAGGGGAGTGGTGGAGG 960
FT	source 1.: 3343 /organism='Microbacterium liquefaciens'.	961	GATGTTGATGAATTTCAAATCTCAGGCCATGGTGGCAGGTCGGCTCCTACCATGGTGGAGGCC 1020
FT	/organism="Microbacterium liquefaciens" /mol_type="genomic DNA" /db_xref="taxon:33918"	3065	GCGGTGAACGGGTTCTCGGCTCAGGCCATGGTGGCTCTGGTGGATCGGC 3124
FEATURES	source	1021	GTTCGGAAAGGGCCCTCGGCCCTACACACCGGGATATCGCAGTGGGGGGCC 1080
ORIGIN	Query Match 59.7%; Score 739.8; DB 6; Length 3343; Best Local Similarity 74.8%; Pred. No. 3.5e-120; Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;	3125	GTGGCGAACGGGGCTCGAACATCGGCTTCAACCGAGGGCATCGG 3184
Db	1 GTGACCCCTGAGAAAGGCCAACGGGGCGCATTGAGAAAGAGATCCGGAGCTCTCCGG 60	1081	CACGACTCGATGTTCATGGCCCATGGGACCTACGGGACCTGGTTC 1140
Db	2105 GTGACGCTGAGGGCAGGGCGGCGATGCCATGGCAGGGAGGCTCTGGACTCTCCCGC 2164	3185	CACGACTCGATGTTCATGGCCCATGGGACCTACGGGACCTGGT 3244
Qy	61 TTCTCGGAGAAAGGCCCGGTTACCCGGCTGACCTACACTCCAGAGCATGCCGGCG 120	1141	GCTGGCCGGAGCCACGTTCCGAAGAATGGGACCTTGACCTTCGCAAAAGGA 1200
Db	2165 TTCTCGGTGAGGGGCCCGGGCTGACACGTCTCACGGTACACTCCGGAGCACGCCGGCG 2224	3245	GCCGGGGAAAGCCATGTGCCGAGGAATGGTCCGATCTGGGAAGGGGAGC 3304
Qy	121 CGGGAAACGGTCATTGGCTATGAAAGGGCCCTTGAGGCTTCGTTGAAGAACGCACTC 180	1201	GAGGTTGTCCTCCGGTAATATGAAGGGCATTGACCC 1239
Db	2225 CGAGGGTGTATCGTCGGCCATGCAAGGGCTGAGGGTCCACGGACGGCTCTC 2284	3305	GATGTTGGCCTCACTGTCGGCTGACGGGGCTGACCC 3343
Qy	181 GGAAACATCATGGCCGACGGCACTGATGCCGAGCTTCTGGATCGGGTCGGT 240	RESULT 15	
Db	2285 GGCAACATCATGGCTGGGTGAGGGGAGGCCCGCTCTGGCGGATCGCCTCGGC 2344	BD181026	3343 bp DNA linear PAT 15-MAY-2003
Qy	241 TCACACTTCGATTCTGTCGGCATGGGACTGTTGATGGCACTGCAAGGGCTACGGTGAATCGGCATCCATTGAG 300	LOCUS	5-Substituted hydantoin racemase, DNA encoding the same, recombinant DNA, transformed cells and process for producing optically active amino acid.
Db	2345 TCGCACTTCGACTCGGTCCGGCAACGGGGCTGTTGACGGGACCGGGCTGGTGC 2404	DEFINITION	BD181026
Qy	301 GCCCTTGAGGCTGCCGGGTGATGGCTGGAGAACGGGCTACGGTGAATCGGCATCCATTGAG 360	VERSION	GI:30791944
Db	2405 GCGCTCGAGGCTGCGAGGGCTGCTGGAGGGATATGTGAACCGTCACTCCTCTCGAG 2464	KEYWORDS	JP 2002330784-A/4.
Qy	361 TTCATCGGAGGGTGGCTGCCGACAGGGAAACGGGGCTTACGGGACTCTGGGAGGATGGGCTGGGGGGCTGG 420	SOURCE	Microbacterium liquefaciens
Db	2465 GTCATCGGAGTCGTGCAAGGGGGCACCGCTTACGGGAGGGATATGTGAACCGTCACTCCTCTCGAG 2524	ORGANISM	Microbacterium liquefaciens Bacteria; Actinobacteria; Actinomycetales; Micrococcineae; Microbacteriaceae; Microbacterium.
Qy	421 GCGATTGCGAGGGTGGCTGCCGACAGGGAAACGGGACTCTGGGAGGATGGGCTGGGGGGCTGGGGGGCTGG 480	REFERENCE	Suzuki, S., Onishi, N. and Yokozeki, K.
Db	2525 GCGATCGGGGCTCGTCCGACGGGGCACCGCTTACGGGAGGGATCTGGACCCCTGTCGGACGGGGCTGG 584	AUTHORS	5-Substituted hydantoin racemase, DNA encoding the same, recombinant DNA, transformed cells and process for producing optically active
Qy	481 TCCGTTAGGCAGGGGGCTACTGCCCTTGGCTTGAAGGCCGGGAACGGCTGGAGGCC 540	TITLE	Patent: JP 2002330784-A 4 19-NOV-2002;
COMMENT	OS Microbacterium liquefaciens	JOURNAL	OS Microbacterium CO INC
PN	JP 2002330784-A/4	COMMENT	OS Microbacterium liquefaciens

PD	1.9-NOV-2002	GGCGTGGTGCCTGGGGGCGCTCATGGTGGAGAGATCAATCGGTTGTCACGGAGATCGGCC	2884
PF	13-SEP-2001	JP 2001278739	
PI	SHUNICHI SUZUKI, NORIMASA ONISHI, KENZO YOKOZEKI PC		
C12N15/09, C12N1/21, C12N9/90, C12P13/02, C12P17/10, PC			
C12N15/00			
CC	5-Substituted hydantoin racemase, DNA encoding the same, CC recombinant DNA, transformed cells and process for producing optically active amino acid		
Key	Location/Qualifiers		
source	1. .3343 /organism='Microbacterium liquefaciens'.		
FT	Location/Qualifiers		
FEATURES	1. .3343 /organism="Microbacterium liquefaciens"		
SOURCE	/mol_type="genomic DNA"		
	/db_xref="taxon:33918"		
ORIGIN			
Query Match	59.7%; Score 739.8; DB 6; Length 3343;		
Best Local Similarity	74.8%; Pred. No. 3.5e-120;		
Matches	927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;		
Qy	1 GTGACCCCTGGAGAAAGGCCAAGGGGGGGCATTTGAGAAAGAGATCCGGAGCTCTCCGG	60	
Db	2105 GTGACCGCTGGAGCAGGGGGGGCATCGGAGGTCTGGACTCTCTCCGG	2164	
Qy	61 TTCTCGGCGAGAAGGGCCCCGGTGTACCCGGCTGACCTACACTCCAGGAGATGCCCG	120	
Db	2165 TTCTCGGTCAAAGGGCCCCGGCGTGAACAGTCTCACGTACACTCCGGAGCACGCCCG	2224	
Qy	121 CGGAAAACCTCAATTGGGCTATGAAAGGGGCCCTTGAGGGTTCTGTAAGACGCACTC	180	
Db	2225 CGAGAGGTGATCGTCGCCCATGCAAGGGGACGGGTGAGGGTCCACGAGGACGCTCTC	2284	
Qy	181 GGAAACATCATTCGGCCGACGTGAAGGCACTGATCCGGAGCTTCCTGGATTCGGT	240	
Db	2285 GGCAACATCATCGGTGGCTGGGTGAGGGGAGGCCCGCTGCGGAGATGGCTCGGC	2344	
Qy	241 TCACACTTCGATTCTGTCCGAAACGGGGGATGTTGATGGCACTGCAGGGTGGTGTGC	300	
Db	2345 TCGCACTTCGACTCGGTCCGCAACGGGGGGATGTTGACGGGGATATGGTCAACGG	2404	
Qy	301 GCCCTTGAGGGCTGCCGGGTGATGGAGAACGGCTACGGCATGGCATTGAG	360	
Db	2405 GCGCTCGAGGGCTCGTGTGAGGGGACCCGGCTGACGGGGCATGGCTCGAG	2404	
Qy	361 TTCAATTGGCTGGGGCTACTGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	420	
Db	2465 GTCAATTGGCTGGGGCTCGTGTGAGGGGACCCGGCTTCAAGCAGGGCATGGCT	2524	
Qy	421 GCCATTGGCAAGGGGGCTACTGGCTTGGGGGGGGGGGGGGGGGGGGGGGGGGGG	480	
Db	2525 GCGATCGGGGGCTCGTGTGAGGGGACCCGGCTTCAAGCAGGGCATGGCTCGAG	2584	
Qy	481 TCCGGTAGGCAAGGGGGACTTGGGTGCTGGAGGGAAACTGGGACTCTTGGTGTAGGGAGTG	540	
Db	2585 ACGGTGCGGGATGACCTTCGGCGCTTCAAGCAGGGGGCATGGCTCGAG	2644	
Qy	541 CGCTCCGGTCAAAGGCCAGAAGCGAACCGCCCTTATCGAAACTACACATTGAAACAGGACCGATGCC	600	
Db	2645 CGTAGGGAGGATGACCTTCGGCGCTTCAAGCAGGGGGATCGTGGACATCGAG	2704	
Qy	601 GAGCAGGAGCAAATAGAGATCGGAGTTGTAACCTCCATCGTGGCTTGGGCAATTGGG	660	
Db	2705 GAGCAGGAGAAGGTGGAGATCGGGGTGCTGGACGGGGATCGTGGCTTCCGG	2764	
Qy	661 GRTGCCGTCAAAGGCCAGAAGCGAACCGCCCTTATGCACCCCCATCGTGGCTTGGGCAATTGGG	720	
Db	2765 ATCACGGTGGGGAGGAGGACCGCCACGGGGACGACCTGGGGATCGTCAACGGAGATGCC	2824	
Qy	721 GCGCTGGTACCCGGCTCTCATGGTGGAGGTCAAACCGGTTCGTCAACGGAGATGCC	780	

This Page Blank (uspto)

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: May 4, 2004, 12:19:37 ; Search time 614 Seconds

(without alignments)
 8572.504 Million cell updates/sec

Title: US-10-045-063-1

Perfect score: 1239
 Sequence: tggaccctgccccaaaggcca.....tgaaggcacccgtaccgttaa 1239Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing First 45 summaries

Database : N_Geneseq_29Jan04:
 1: _geneseqn1980s:
 2: _geneseqn1990s:
 3: _geneseqn2000s:
 4: _geneseqn2001as:
 5: _geneseqn2001bs:
 6: _geneseqn2002as:
 7: _geneseqn2003as:
 8: _geneseqn2003bs:
 9: _geneseqn2003cs:
 10: _geneseqn2004s:
 11: _geneseqn2005as:
 12: _geneseqn2005bs:
 13: _geneseqn2005cs:
 14: _geneseqn2006as:
 15: _geneseqn2006bs:
 16: _geneseqn2006cs:
 17: _geneseqn2007as:
 18: _geneseqn2007bs:
 19: _geneseqn2007cs:
 20: _geneseqn2008as:
 21: _geneseqn2008bs:
 22: _geneseqn2008cs:
 23: _geneseqn2009as:
 24: _geneseqn2009bs:
 25: _geneseqn2009cs:
 26: _geneseqn2010as:
 27: _geneseqn2010bs:
 28: _geneseqn2010cs:
 29: _geneseqn2011as:
 30: _geneseqn2011bs:
 31: _geneseqn2011cs:
 32: _geneseqn2012as:
 33: _geneseqn2012bs:
 34: _geneseqn2012cs:
 35: _geneseqn2013as:
 36: _geneseqn2013bs:
 37: _geneseqn2013cs:
 38: _geneseqn2014as:
 39: _geneseqn2014bs:
 40: _geneseqn2014cs:
 41: _geneseqn2015as:
 42: _geneseqn2015bs:
 43: _geneseqn2015cs:
 44: _geneseqn2016as:
 45: _geneseqn2016bs:
 46: _geneseqn2016cs:
 47: _geneseqn2017as:
 48: _geneseqn2017bs:
 49: _geneseqn2017cs:
 50: _geneseqn2018as:
 51: _geneseqn2018bs:
 52: _geneseqn2018cs:
 53: _geneseqn2019as:
 54: _geneseqn2019bs:
 55: _geneseqn2019cs:
 56: _geneseqn2020as:
 57: _geneseqn2020bs:
 58: _geneseqn2020cs:
 59: _geneseqn2021as:
 60: _geneseqn2021bs:
 61: _geneseqn2021cs:
 62: _geneseqn2022as:
 63: _geneseqn2022bs:
 64: _geneseqn2022cs:
 65: _geneseqn2023as:
 66: _geneseqn2023bs:
 67: _geneseqn2023cs:
 68: _geneseqn2024as:
 69: _geneseqn2024bs:
 70: _geneseqn2024cs:
 71: _geneseqn2025as:
 72: _geneseqn2025bs:
 73: _geneseqn2025cs:
 74: _geneseqn2026as:
 75: _geneseqn2026bs:
 76: _geneseqn2026cs:
 77: _geneseqn2027as:
 78: _geneseqn2027bs:
 79: _geneseqn2027cs:
 80: _geneseqn2028as:
 81: _geneseqn2028bs:
 82: _geneseqn2028cs:
 83: _geneseqn2029as:
 84: _geneseqn2029bs:
 85: _geneseqn2029cs:
 86: _geneseqn2030as:
 87: _geneseqn2030bs:
 88: _geneseqn2030cs:
 89: _geneseqn2031as:
 90: _geneseqn2031bs:
 91: _geneseqn2031cs:
 92: _geneseqn2032as:
 93: _geneseqn2032bs:
 94: _geneseqn2032cs:
 95: _geneseqn2033as:
 96: _geneseqn2033bs:
 97: _geneseqn2033cs:
 98: _geneseqn2034as:
 99: _geneseqn2034bs:
 100: _geneseqn2034cs:
 101: _geneseqn2035as:
 102: _geneseqn2035bs:
 103: _geneseqn2035cs:
 104: _geneseqn2036as:
 105: _geneseqn2036bs:
 106: _geneseqn2036cs:
 107: _geneseqn2037as:
 108: _geneseqn2037bs:
 109: _geneseqn2037cs:
 110: _geneseqn2038as:
 111: _geneseqn2038bs:
 112: _geneseqn2038cs:
 113: _geneseqn2039as:
 114: _geneseqn2039bs:
 115: _geneseqn2039cs:
 116: _geneseqn2040as:
 117: _geneseqn2040bs:
 118: _geneseqn2040cs:
 119: _geneseqn2041as:
 120: _geneseqn2041bs:
 121: _geneseqn2041cs:
 122: _geneseqn2042as:
 123: _geneseqn2042bs:
 124: _geneseqn2042cs:
 125: _geneseqn2043as:
 126: _geneseqn2043bs:
 127: _geneseqn2043cs:
 128: _geneseqn2044as:
 129: _geneseqn2044bs:
 130: _geneseqn2044cs:
 131: _geneseqn2045as:
 132: _geneseqn2045bs:
 133: _geneseqn2045cs:
 134: _geneseqn2046as:
 135: _geneseqn2046bs:
 136: _geneseqn2046cs:
 137: _geneseqn2047as:
 138: _geneseqn2047bs:
 139: _geneseqn2047cs:
 140: _geneseqn2048as:
 141: _geneseqn2048bs:
 142: _geneseqn2048cs:
 143: _geneseqn2049as:
 144: _geneseqn2049bs:
 145: _geneseqn2049cs:
 146: _geneseqn2050as:
 147: _geneseqn2050bs:
 148: _geneseqn2050cs:
 149: _geneseqn2051as:
 150: _geneseqn2051bs:
 151: _geneseqn2051cs:
 152: _geneseqn2052as:
 153: _geneseqn2052bs:
 154: _geneseqn2052cs:
 155: _geneseqn2053as:
 156: _geneseqn2053bs:
 157: _geneseqn2053cs:
 158: _geneseqn2054as:
 159: _geneseqn2054bs:
 160: _geneseqn2054cs:
 161: _geneseqn2055as:
 162: _geneseqn2055bs:
 163: _geneseqn2055cs:
 164: _geneseqn2056as:
 165: _geneseqn2056bs:
 166: _geneseqn2056cs:
 167: _geneseqn2057as:
 168: _geneseqn2057bs:
 169: _geneseqn2057cs:
 170: _geneseqn2058as:
 171: _geneseqn2058bs:
 172: _geneseqn2058cs:
 173: _geneseqn2059as:
 174: _geneseqn2059bs:
 175: _geneseqn2059cs:
 176: _geneseqn2060as:
 177: _geneseqn2060bs:
 178: _geneseqn2060cs:
 179: _geneseqn2061as:
 180: _geneseqn2061bs:
 181: _geneseqn2061cs:
 182: _geneseqn2062as:
 183: _geneseqn2062bs:
 184: _geneseqn2062cs:
 185: _geneseqn2063as:
 186: _geneseqn2063bs:
 187: _geneseqn2063cs:
 188: _geneseqn2064as:
 189: _geneseqn2064bs:
 190: _geneseqn2064cs:
 191: _geneseqn2065as:
 192: _geneseqn2065bs:
 193: _geneseqn2065cs:
 194: _geneseqn2066as:
 195: _geneseqn2066bs:
 196: _geneseqn2066cs:
 197: _geneseqn2067as:
 198: _geneseqn2067bs:
 199: _geneseqn2067cs:
 200: _geneseqn2068as:
 201: _geneseqn2068bs:
 202: _geneseqn2068cs:
 203: _geneseqn2069as:
 204: _geneseqn2069bs:
 205: _geneseqn2069cs:
 206: _geneseqn2070as:
 207: _geneseqn2070bs:
 208: _geneseqn2070cs:
 209: _geneseqn2071as:
 210: _geneseqn2071bs:
 211: _geneseqn2071cs:
 212: _geneseqn2072as:
 213: _geneseqn2072bs:
 214: _geneseqn2072cs:
 215: _geneseqn2073as:
 216: _geneseqn2073bs:
 217: _geneseqn2073cs:
 218: _geneseqn2074as:
 219: _geneseqn2074bs:
 220: _geneseqn2074cs:
 221: _geneseqn2075as:
 222: _geneseqn2075bs:
 223: _geneseqn2075cs:
 224: _geneseqn2076as:
 225: _geneseqn2076bs:
 226: _geneseqn2076cs:
 227: _geneseqn2077as:
 228: _geneseqn2077bs:
 229: _geneseqn2077cs:
 230: _geneseqn2078as:
 231: _geneseqn2078bs:
 232: _geneseqn2078cs:
 233: _geneseqn2079as:
 234: _geneseqn2079bs:
 235: _geneseqn2079cs:
 236: _geneseqn2080as:
 237: _geneseqn2080bs:
 238: _geneseqn2080cs:
 239: _geneseqn2081as:
 240: _geneseqn2081bs:
 241: _geneseqn2081cs:
 242: _geneseqn2082as:
 243: _geneseqn2082bs:
 244: _geneseqn2082cs:
 245: _geneseqn2083as:
 246: _geneseqn2083bs:
 247: _geneseqn2083cs:
 248: _geneseqn2084as:
 249: _geneseqn2084bs:
 250: _geneseqn2084cs:
 251: _geneseqn2085as:
 252: _geneseqn2085bs:
 253: _geneseqn2085cs:
 254: _geneseqn2086as:
 255: _geneseqn2086bs:
 256: _geneseqn2086cs:
 257: _geneseqn2087as:
 258: _geneseqn2087bs:
 259: _geneseqn2087cs:
 260: _geneseqn2088as:
 261: _geneseqn2088bs:
 262: _geneseqn2088cs:
 263: _geneseqn2089as:
 264: _geneseqn2089bs:
 265: _geneseqn2089cs:
 266: _geneseqn2090as:
 267: _geneseqn2090bs:
 268: _geneseqn2090cs:
 269: _geneseqn2091as:
 270: _geneseqn2091bs:
 271: _geneseqn2091cs:
 272: _geneseqn2092as:
 273: _geneseqn2092bs:
 274: _geneseqn2092cs:
 275: _geneseqn2093as:
 276: _geneseqn2093bs:
 277: _geneseqn2093cs:
 278: _geneseqn2094as:
 279: _geneseqn2094bs:
 280: _geneseqn2094cs:
 281: _geneseqn2095as:
 282: _geneseqn2095bs:
 283: _geneseqn2095cs:
 284: _geneseqn2096as:
 285: _geneseqn2096bs:
 286: _geneseqn2096cs:
 287: _geneseqn2097as:
 288: _geneseqn2097bs:
 289: _geneseqn2097cs:
 290: _geneseqn2098as:
 291: _geneseqn2098bs:
 292: _geneseqn2098cs:
 293: _geneseqn2099as:
 294: _geneseqn2099bs:
 295: _geneseqn2099cs:
 296: _geneseqn2100as:
 297: _geneseqn2100bs:
 298: _geneseqn2100cs:
 299: _geneseqn2101as:
 300: _geneseqn2101bs:
 301: _geneseqn2101cs:
 302: _geneseqn2102as:
 303: _geneseqn2102bs:
 304: _geneseqn2102cs:
 305: _geneseqn2103as:
 306: _geneseqn2103bs:
 307: _geneseqn2103cs:
 308: _geneseqn2104as:
 309: _geneseqn2104bs:
 310: _geneseqn2104cs:
 311: _geneseqn2105as:
 312: _geneseqn2105bs:
 313: _geneseqn2105cs:
 314: _geneseqn2106as:
 315: _geneseqn2106bs:
 316: _geneseqn2106cs:
 317: _geneseqn2107as:
 318: _geneseqn2107bs:
 319: _geneseqn2107cs:
 320: _geneseqn2108as:
 321: _geneseqn2108bs:
 322: _geneseqn2108cs:
 323: _geneseqn2109as:
 324: _geneseqn2109bs:
 325: _geneseqn2109cs:
 326: _geneseqn2110as:
 327: _geneseqn2110bs:
 328: _geneseqn2110cs:
 329: _geneseqn2111as:
 330: _geneseqn2111bs:
 331: _geneseqn2111cs:
 332: _geneseqn2112as:
 333: _geneseqn2112bs:
 334: _geneseqn2112cs:
 335: _geneseqn2113as:
 336: _geneseqn2113bs:
 337: _geneseqn2113cs:
 338: _geneseqn2114as:
 339: _geneseqn2114bs:
 340: _geneseqn2114cs:
 341: _geneseqn2115as:
 342: _geneseqn2115bs:
 343: _geneseqn2115cs:
 344: _geneseqn2116as:
 345: _geneseqn2116bs:
 346: _geneseqn2116cs:
 347: _geneseqn2117as:
 348: _geneseqn2117bs:
 349: _geneseqn2117cs:
 350: _geneseqn2118as:
 351: _geneseqn2118bs:
 352: _geneseqn2118cs:
 353: _geneseqn2119as:
 354: _geneseqn2119bs:
 355: _geneseqn2119cs:
 356: _geneseqn2120as:
 357: _geneseqn2120bs:
 358: _geneseqn2120cs:
 359: _geneseqn2121as:
 360: _geneseqn2121bs:
 361: _geneseqn2121cs:
 362: _geneseqn2122as:
 363: _geneseqn2122bs:
 364: _geneseqn2122cs:
 365: _geneseqn2123as:
 366: _geneseqn2123bs:
 367: _geneseqn2123cs:
 368: _geneseqn2124as:
 369: _geneseqn2124bs:
 370: _geneseqn2124cs:
 371: _geneseqn2125as:
 372: _geneseqn2125bs:
 373: _geneseqn2125cs:
 374: _geneseqn2126as:
 375: _geneseqn2126bs:
 376: _geneseqn2126cs:
 377: _geneseqn2127as:
 378: _geneseqn2127bs:
 379: _geneseqn2127cs:
 380: _geneseqn2128as:
 381: _geneseqn2128bs:
 382: _geneseqn2128cs:
 383: _geneseqn2129as:
 384: _geneseqn2129bs:
 385: _geneseqn2129cs:
 386: _geneseqn2130as:
 387: _geneseqn2130bs:
 388: _geneseqn2130cs:
 389: _geneseqn2131as:
 390: _geneseqn2131bs:
 391: _geneseqn2131cs:
 392: _geneseqn2132as:
 393

CC DNA. A whole cell catalyst for the conversion of 5-monosubstituted hydantoins to L- or D-amino acids, comprises cloned genes encoding for a hydantoinase, a racemase and a carbamoylase. These genes are overexpressed in the cell according to their turnover rates. The bioconversion consists of chemical and/or enzymatic racemisation of 5-substituted hydantoins, ring opening hydrolysis achieved by a hydantoinase, and hydrolysis of the N-carbamoyl amino acid produced by hydantoinase to the amino acid by carbamoylase. Isolation of microorganisms capable for stereospecific hydrolysis of hydantoins is necessary due to the increasing demand for optically pure amino acids. The cells are usually bacterial cells, e.g. from Escherichia coli, since they have high reproduction rates and easy growing conditions. Since all the enzymes are produced in one strain, there is a reduction of fermentation and purification costs.

XX Sequence 1239 BP; 234 A; 357 C; 406 G; 242 T; 0 U; 0 Other;

Query Match 99.9%; Score 1238; DB 5; Length 1239;

Best Local Similarity 100.0%; Pred. No. 1.1e-287; Mismatches 0; Indels 0; Gaps 0;

Matches 1238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 TGACCCCTGGCAGAAAGCGCAAGCGGGCGCATTGAGAAAAGAGATCCGGGGAGGTCTCCCGGT 61

Db 2 TGACCCCTGGCAGAAAGCGCAAGCGGGCGCATTGAGAAAAGAGATCCGGGGAGGTCTCCCGGT 61

Qy 62 TCTCGGGCAGAAGGCCCGGGTTACCCGGCTGTACCTACACTCCAGAGCATGCCGC 121

Db 62 TCTCGGGCAGAAGGCCCGGGTTACCCGGCTGTACCTACACTCCAGAGCATGCCGC 121

Qy 122 GGGAAACGGCTCATGGGGCTATGAAAGCGGGCCTTGAGCGTTGCTGAAGAACGGCACTCG 181

Db 122 GGGAAACGGCTCATGGGGCTATGAAAGCGGGCCTTGAGCGTTGCTGAAGAACGGCACTCG 181

Qy 182 GAAACATCATGGGGCAGCTGAGGCACTGATCCGGAGCTTCCTGGCATGGGGT 241

Db 182 GAAACATCATGGGGCAGCTGAGGCACTGATCCGGAGCTTCCTGGCATGGGGT 241

Qy 242 CACACTCGATTCTCGGAACAGGGGGACTGATCGGGATCGGGCTTCCTGGCATGGGGT 301

Db 242 CACACTCGATTCTCGGAACAGGGGGACTGATCGGGCTTCCTGGCATGGGGT 301

Qy 302 CCCTTGAGGGTGCCTGGATGCTGGAAACGGGTACGTGAATGGCATTCATTGAGT 361

Db 302 CCCTTGAGGGTGCCTGGATGCTGGAAACGGGTACGTGAATGGCATTCATTGAGT 361

Qy 362 TCATCGGATCGTGGAGGAAGGGGAACAGGGAACTGGGACTCTTGGTGTAGGGATGGAGT 481

Db 362 TCATCGGATCGTGGAGGAAGGGGAACAGGGAACTGGGACTCTTGGTGTAGGGATGGAGT 481

Qy 422 CCATTGACGGGTTGGTGCCTGACTGGGACTCTTGGTGTAGGGATGGAGT 481

Db 422 CCATTGACGGGTTGGTGCCTGACTGGGACTCTTGGTGTAGGGATGGAGT 481

Qy 482 CGCTTAAGGCCAGGGCTACTGCCCTTGGCTGAAGGGGGGAACGGGTGCAGGCC 541

Db 482 CGCTTAAGGCCAGGGCTACTGCCCTTGGCTGAAGGGGGGAACGGGTGCAGGCC 541

Qy 542 GCTCCGGGGGACCTGGTGCCTGACTGGGACTCTGGTGTAGGGATGGAGT 481

Db 542 GCTCCGGGGGACCTGGTGCCTGACTGGGACTCTGGTGTAGGGATGGAGT 481

Qy 602 AGCAGGAGCAAATAGAGATCGGAGTTGTAACCTCCATGGAAACTACACATTGAA 661

Db 602 AGCAGGAGCAAATAGAGATCGGAGTTGTAACCTCCATGGAAACTACACATTGAA 661

Qy 662 TTGCCGTCAAAGGAGAACACGGGACCAACCCCCATGCCACCTGGCCAGGATG 721

Db 662 TTGCCGTCAAAGGAGAACACGGGACCAACCCCCATGCCACCTGGCCAGGATG 721

Qy 722 CGCTGGTACCCGGCGCTCTCATGGTGGAGGTCAACGGGTTCGTCAAGAGATCGCG 781

Db 722 CGCTGGTACCCGGCGCTCTCATGGTGGAGGTCAACGGGTTCGTCAAGAGATCGCG 781

CC	Qy 782 ATGGCACAGTGGCTACCGTTGGCACCTCACAGTGGCCCCGGTGGAGGCAACCAAGGTCC 841	Db 782 ATGGCACAGTGGCTACCGTTGGCACCTCACAGTGGCCCCGGTGGAGGCAACCAAGGTCC 841
CC	Qy 842 CGGGGGAGGTGGACTTCAACTGACCTGCGTTCTCGCATGGAGTCGCTCCGGTGC 901	Db 842 CGGGGGAGGTGGACTTCAACTGACCTGCGTTCTCGCATGGAGTCGCTCCGGTGC 901
CC	Qy 902 TGATCGACCGCATCTGGTCACTGGTGCAGGAGTCGCTCCGGTGC 961	Db 902 TGATCGACCGCATCTGGTCACTGGTGCAGGAGTCGCTCCGGTGC 961
CC	Qy 962 ATGGGGATGAATTTCATCTCAGCCCCGGTGCAGCTGGCTCTACCATGGTGGACGCC 1021	Db 962 ATGTGGGATGAATTTCATCTCAGCCCCGGTGCAGCTGGCTCTACCATGGTGGACGCC 1021
SQ	Sequence 1239 BP; 234 A; 357 C; 406 G; 242 T; 0 U; 0 Other;	
XX	Query Match 99.9%; Score 1238; DB 5; Length 1239;	Qy 1022 TTCGGAAAGGGCCTCGGCCCTCGAGTTCACACACGGGATATCAGCAGTGGGGGGCC 1081
XX	Best Local Similarity 100.0%; Pred. No. 1.1e-287;	Db 1022 TCGGAAAGGGCCTCGGCCCTCGAGTTCACACACGGGATATCAGCAGTGGGGGGCC 1081
XX	Matches 1238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	Qy 1082 ACGACTCGATGGTTCATGCCCAAGGTACGGAGCTGGAAATGGTTTCAGGCCGTG 1141
XX		Db 1082 ACGACTCGATGGTTCATGCCCAAGGTACGGAGCTGGAAATGGTTTCAGGCCGTG 1141
XX		Qy 1142 CTGGCCGGAGCCACGTCCCGAAGAATGGACCCGATTCGACCTTCGCAAGGAAACTG 1201
XX		Db 1142 CTGGCCGGAGCCACGTCCCGAAGAATGGACCCGATTCGACCTTCGCAAGGAAACTG 1201
XX		Qy 1202 AGGTTGTCTCCGGTAATGAGGGCACTTGAACCGATTGACCTTGAACCGGTAA 1239
XX		Db 1202 AGGTTGTCTCCGGTAATGAGGGCACTTGAACCGGTAA 1239
XX		RESULT 2
XX		ID ABL41208 standard; DNA; 1239 BP.
XX		XX AC ABL41208;
XX		DT 07-MAY-2002 (First entry)
XX		XX DE Arthrobacter aurescens polynucleotide SEQ ID NO 3.
XX		XX KW Allysine acetal; hydantoin; carbamoylase; racemisation; pharmaceutical; ds.
XX		XX OS Arthrobacter aurescens.
XX		PN DE10037115-A1.
XX		PD 07-FEB-2002.
XX		PF 28-JUL-2000; 20000DE-01037115.
XX		PR 28-JUL-2000; 20000DE-01037115.
XX		XX (DEGS) DEGUSSA AG.
XX		PA Krimmer H, May O, Klement I, Drauz K, Reichert D;
XX		DR WPI; 2002-218024/28.
XX		PS Disclosure; Page 5; 6pp; German.
CC	The invention relates to the preparation of allysine acetals (I) from corresponding hydantoins (II) involving reaction of (II) with spontaneous hydantoinase and D- or L-specific carbamoylase, together with spontaneous and/or enzyme-catalysed racemisation, where the enzymes are used in free, immobilised or cell-enclosed form. The use of (I) obtained by the process	CC

is claimed in the synthesis of biologically active agents, especially pharmaceuticals. In particular (I) are intermediates for pharmaceuticals as described in US5552397, WO9738705 and J. Med. Chem. 42, 305 (1999). The process is easy to carry out and suitable for large-scale use. The acetal protecting group is stable under the reaction conditions and (I) can be obtained in very high yield. The L-isomers of (I) (generally the isomers required for drug synthesis) are readily obtainable by using L-carbamoylase. The present sequence is that of an *Arthrobacter aurescens* polynucleotide useful to the invention.

is claimed in the synthesis of biologically active agents, especially pharmaceuticals. In particular (I) are intermediates for pharmaceuticals as described in US5552397, WO9738705 and J. Med. Chem. 42, 305 (1999). The process is easy to carry out and suitable for large-scale use. The acetal protecting group is stable under the reaction conditions and (I) can be obtained in very high yield. The L-isomers of (I) (generally the isomers required for drug synthesis) are readily obtainable by using L-carbamoylase. The present sequence is that of an Arthrobacter aurescens polynucleotide useful to the invention

Sequence 1239 BP; 234 A; 357 C; 406 G; 242 T; 0 U; 0 Other;

Query Match 99.9%; Score 1238; DB 6; Length 1239;
Best Local Similarity 100.0%; Pred. No. 1.1e-287;
Matches 1238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

b	2	TGACCCCTGCAGAAAGCGCAAGCGGCCATTGAGAAAGAGATCCGGAGCTCTCCGGT	61
b	2	TGACCCGTGAGAAAGGCCAACGGCCAGCTACACTCCAGAGCATGCCCGC	61
b	62	TCTGGCAGAAGGCCCCGGTGTACCCGGCTGACCTACACTCCAGAGCATGCCCGC	121
b	62	TCTGGCAGAAGGCCCCGGTGTACCCGGCTGACCTACACTCCAGAGCATGCCCGC	121
b	122	GGGAAACGGCTCATTGGGCTATGAAAGCCGCCCTTGAGCGTTCTGTAAGAACGCACTCG	181
b	122	GGGAAACGGCTCATTGGGCTATGCGGCTATGAAAGCCGCCCTTGAGCGTTCTGTAAGAACGCACTCG	181
b	182	GAAACATCATCGGCCGACGTGAAGGCACTGATCCGGAGCTTCCTGCGATCGGGT	241
b	182	GAAACATCATCGGCCGACGTGAAGGCACTGATCCGGAGCTTCCTGCGATCGGGT	241
b	242	CACACTTCGATTCTGTCGGAAACGGCGGATGTTGATGGCACTGCAAGGGTGGTGG	301
b	242	CACACTTCGATTCTGTCGGCACTGTAAGGCACTGATCCGGAGCTTCCTGCGATCGGGT	241
b	302	CCCTTGAGGTGCCGGTGTATGGGAGAACGGGCTACGTGAATCGGCAATCCATTGAGT	361
b	302	CCCTTGAGGTGCCGGTGTATGGGAGAACGGGCTACGTGAATCGGCAATCCATTGAGT	361
b	362	TCATCGCAATCGTGGAGAACGGGCTACGTGAATCGGCAATCCATTGAGT	361
b	362	TCATCGCAATCGTGGAGAACGGGCTACGTGAATCGGCAATCCATTGAGT	361
b	362	TCATCGCAATCGTGGAGAACGGGCTACGTGAATCGGCAATCCATTGAGT	361
b	422	CCATTGCAAGGGTTGGTGCCTGGAGAACGGGAACCTGGACTCTTGGTTGATGAGGATGGAGTGT	421
b	422	CCATTGCAAGGGTTGGTGCCTGGAGAACGGGAACCTGGACTCTTGGTTGATGAGGATGGAGTGT	421
b	482	CCGTTAGGCAAGGGCTACTGCCTTCGGCTTGAGGGCAACTGCAGGCTGCAGCCC	541
b	482	CCGTTAGGCAAGGGCTACTGCCTTCGGCTTGAGGGCAACTGCAGGCTGCAGCCC	541
b	542	GCTCCGGCGGGGACCTGGCTGCTTTATCGAACTACATCGAAACAGGACCGATCCTCG	601
b	542	GCTCCGGCGGGGACCTGGCTGCTTTATCGAACTACATCGAAACAGGACCGATCCTCG	601
b	602	AGCAGGAGCAAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCTGCGCATCGGG	661
b	602	AGCAGGAGCAAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCTGCGCATCGGG	661
b	662	TGCGGTCAAAGGCAGAACGGCAACCGGGCACACCCCATGCACCTGCGCAGGATG	721
b	662	TGCGGTCAAAGGCAGAACGGCAACCGGGCACACCCCATGCACCTGCGCAGGATG	721
b	722	CGCTGGTACCCGGCTCTCATGGTGGAGGTCAACCGGTTCTGTCACAGAGATCGCG	781
b	722	CGCTGGTACCCGGCTCTCATGGTGGAGGTCAACCGGTTCTGTCACAGAGATCGCG	781
b	782	ATGGCACAGGGCTACCGTTGGCCACCTCACAGTGGCCCCGGTGGAGGCAACCGGT	841
b	782	ATGGCACAGGGCTACCGTTGGCTACGGAGTGGCTCCGGCATAGGAGTGGCTCCGGT	901

RESULT 3
ABV72490
ID ABV72490 standard; DNA; 1239 BP.
AC ABV72490;
XX DT 29-JAN-2003 (first entry)
XX DE Nucleotide sequence of N-carbamyl-L-amino acid hydrolase enzyme.
XX KW Hydantoinase; enzyme; N-carbamyl-L-amino acid hydrolase;
XX KW optically-active amino acid; gene; ss.
XX OS Microbacterium liquefaciens.
XX FH Location/Qualifiers
XX CDS 1 .1239
FT /*tag= a
FT /product= "hydantoinase"
FT /note= "no start codon given"
XX Key
XX PD 19-SEP-2002.
XX PR 08-MAR-2001; 2001JP-00065814.
XX PR 27-SEP-2001; 2001JP-00298619.
XX PA (AJIN) AJINOMOTO CO INC.
XX PI Takenaka Y, Suzuki S, Onishi N, Yokozeki K;
XX DR 2002-691806/74.
XX DR P-PSDB; ABB99394.
XX Microbacterium-originated DNAs encoding hydantoinase and N-carbamyl-L-amino-acid hydrolase and transformant cells, useful for production of proteins and optically-active amino acids, applicable in synthesis of e.g. pharmaceuticals.
XX PS Claim 13; Page 63-67; 83pp; Japanese.
XX CC The present sequence encodes a N-carbamyl-L-amino acid hydrolase enzyme.

	Sequence	1239 BP; 192 A; 387 C; 462 G; 198 T; 0 U; 0 Other;
Query	Match	Score 59.7%; DB 6; Length 1239; Best Local Similarity 74.8%; Pred. No. 6.1e-168; Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0
b	y	1 GTGACCCCTGGAGAAAGCGCAAGCGGGCATTGAGAAAGAGATCCGGAGCTCTCCGG 60 1 GTGACCGCTGGAGGCGCATCGAGGAGGCTCTGGACTCTTCCGGC 60
b	y	61 TTCTCGGCAGAAGGGCCCCGGTGTACCCGGCTGACCTACACTCCAGAGCATGCCGCCGG 12 61 TTCTCGGTGAAGGGCCCCTCACGTCTCACGGTACACTCCGGAGCACGGCTCTCG 12
b	y	121 CGGGAAACGCTCATTCGGGCTATGAAAGCGGCCCTTGAGCGTTCCGTGAAGACGGCACTC 18 121 CGAGGGTGAATCGTCGCCATGCAAGGGACGGGGCTGAGCGTCCACGAGGACGGCTCTC 18
b	y	181 GGAAACATCATCGGCCGACGGCACTGATCGGAGCTTCCTGGATCGGGATCGGGT 24 181 GGCAACATCATCGGTCTGGGCTGAGGGAGGCCGGCTGCCGGATGCCCTTCGGC 24
b	y	241 TCACACTTCGATTCTGTCCGAAACGGCGGATGTTGATGGCACTGCGAGGGTGTGC 30 241 TCGCACTTCGACTCGGTCCGCAACGGGGATGTTCGACGGCACCGGGGTGGTGTGC 30
b	y	301 GCCCTTGAGGCTGCCCCGGTGTGGAGAACGGCTACGGTAATGGCATCATTGAG 36 301 GCGCTCGAGGGCTGCAAGGGGACCCCGCTTCAGCAGGGCATGTGAACCGTCATCCTCTCGAG 36
b	y	361 TTCAATCGGGATCGTGGAGGAAGGGCCCGCTTCAGCAGTGGCATGTTGGGGCGG 42 361 GTCATCGGGATCGTCGAAGGGGACCCCGCTTCAGCAGGGCATGCTGGGGCATGCTGGGG 42
b	y	421 GCCATTAGGCAGGGCTACTGCCTTCGGCTTGAGGGAAACTGGACTCTTGGTTGATGAGGATGGAGTGTG 48 421 GCGATCGGGCTCGTGTGGACGGCGATCTGGACACCCCTGGTGGACAGGAAGACGGCGTGC 48
b	y	481 TCCGTTAGGCAGGGCTACTGCCTTCGGCTTGAGGGGGGAACACTGGCAGGGCTGCAGGCC 54 481 ACGGGTGGGGCCACGGGCTCGGGCTTGAGGGGTGAGGCTGGGAGGGGG 54
b	y	541 CGCTCCGGGGGACCTGGCTGGTTATCGAACTACACATGAAACAAGGACCGATCCTC 60 541 CGTACGAGGGATGACCTTCGGGCTTCATCGAGTTGCACATCGAGGGCATGGGAGTCCTC 60
b	y	601 GAGCAGGAGCAAATAGAGATCGGAGTTGTAAACCTCCATCGTGGCATTGGGG 66 601 GAGCAGGAGGAAGGTGGAGATCGGGGGATCGTGTGACGGGGATCGTGGGATCGTGGGG 66
b	y	661 GTTGGCGTCAAAGGCAGAAGCGACCAAGCGGGCACACCCCATGCCACCTGGGGCAGGGAT 72 661 ATCACGGTGGGGGAGGAGGACGGGGGACGGGGGACGACCCCATGCACTGGGGCAGGGAC 72
b	y	721 GCGCTGGTACCCGGCTCTCATGGTGAGGGTCAACCGGGCACACCCCATGCCACCTGGGGCAGGGATCGGCC 78 721 GCGCTGGTGGGGGGCTCATGGTGAGGAGATCAATCGGGTGTCAACGGAGATCGGG 78
b	y	781 GATGGCACAGTGGCTTACCGGTTGGCCACCTCACAGTGGCCCCGGTGGAGGCAACCGGGTC 84 781 GACGGCACGGTGGGACCGGACCCACCTCACGGTGATCTGGGCTCAACCAGGGTT 84
b	y	841 CGGGGGAGGTGGACCTTCACACTGGACCTGGCTTCCGCATGAGGAGTCGGCTCCGGGGTGTGGCTGCC 90 841 CCCGGGGCGGTCACTGGGATCTGGCTCACGGGATTCACGGGATTCACGGGCTCAACCAGGGTC 90
b	y	901 CTGATCGACCGCATCTCGGTCAATGGTGGGGAGGTGGCTGCC 96

Db	901	CTGGTCGACAGGGATCGAGGCCATGGTGGCAGGGCCAGAAGTGGCCGGAGTCGAGGGCC	960
QY	961	GATGTGGATGAATTTCATAATCTCAGCCCCGGTGCAGCTGGCTCCTACCATGGTGGACGCC	1020
	961	GGGGTGAAACGGGTTCTCGGCTCAGCCCTGGTCTCGGTCTGGATCGC	1020
QY	1021	GTTCGGCAAGGGCCTGGCCTTGGAGTTCACACAACGGGATATCAGCAGTGGGGGGC	1080
Db	1021	GTGGCGAACGGGTTCCGAACTCGGCTTACCCATCGGACATCACGAGCGGGCAGGG	1080
QY	1081	CACGACTCGATGTTCATGCCAGGTACCGGACGGTCAACGGGATTGGAAATGGTTTCGTTCCAAAGCCGT	1140
Db	1081	CACGACTCGATGTTCATGCCAGATCACCGACGTTGGAAATGGTGTTCGTTCCCAAGCCG	1140
QY	1141	GCTGGCGGAGGCCACGGTTCGGAAAGAATGGGACCGATTTCGATGACCTTCGCAAAGGAAC	1200
Db	1141	GCGGGGCGAAGCCATGGTGGCGGAGGAATGGTCCGGATTTCGACGATCTGGGAAAGGGGACG	1200
QY	1201	GAGGGTTGTCCTCCGGGTAATGAAGGGACTTGACCCGGTAA	1239
Db	1201	GATGTGGTCCTTCACGTCGTTGACGGCGCTTGACCCGGTGA	1239
RESULT 4			
	ABN86382	ABN86382 standard; DNA; 1239 BP.	
	ID	XX	
	XX	ABN86382;	
	XX	DT 21-OCT-2002 (first entry)	
	XX	DE M. liquefaciens Chase encoding DNA.	
	XX	KW Hydantoin racemase; HRase; bacterium; hydantoin; N-carbamoyl-amino acid; pharmaceutical; chemical; food additive; hydantoinase; CHase; gene; ds.	
	XX	OS Microbacterium liquefaciens.	
	XX	KEY Location/Qualifiers	
	XX	CFH CDS 1. .1380	
	FT	/*tag= a	
	FT	/product= "CHase"	
	FT	/note= "N-carbamoyl-amino acid-L-hydrolase"	
	XX	EP1188826-A2.	
	XX	PD 20-MAR-2002.	
	XX	PF 09-AUG-2001; 2001EP-00119275.	
	XX	PR 13-SEP-2000; 2000JP-00278571.	
	PR	08-MAR-2001; 2001JP-00065815.	
	XX	(AJIN) AJINOMOTO CO INC.	
	XX	PI Suzuki S, Onishi N, Yokozeki K;	
	XX	DR WPI; 2002-510588/55.	
	DR	P-PSDB; ABB80945.	
	XX	PT New 5-substituted hydantoin racemase, useful in production of optically active amino acids, comprises high working temperature, from Microbacterium species.	
	XX	PT Disclosure; Page 27-28; 40pp; English.	
	PS	XX The invention relates to a 5-substituted hydantoin racemase (HRase) that is derived from a Microbacterium by culture, disruption or lysis, and purification. HRase has a high working temperature of 5-60 plusoC and is used to racemize optically active 5-substituted hydantoins for subsequent enzymatic conversion to N-carbamoyl-amino acids and then optically active amino acids (useful in pharmaceuticals, the chemical industry and as food additives). The present sequence represents a DNA encoding the M-	

61	TTCTCGGAGAAGGCCCTGGTACCCGGTGTGACTTACACTCCAGAGCATGCCGCCGCC	120	Db	3245 GCGGGCGAACGGCAAGCCATGTGCCGGTAATGAAAGGCACCTTGACGGACCG 3300
2165	TTCTCGGTCAAAGGGCCCCGGTGTGACACGTCTCACGTACACTCCGGAGCAGCGCTCTCGG 2224		Qy	1201 GAGGTGTCCTCCGGTAATGAAAGGCACCTTGACGGACCG 1239
121	CGGAAACACGGCTCATTGGGCTATGAAAGGGCCCTTGAGCGTTCTGAAAGACGCACTC	180	Db	3305 GATGTGTCCTCACGTCTGACGGCTCCACGGAGCACGCTCTCG 3343
2225	CGAGAGGTGATCGTGCAGGCCATGCAAGGGAGCAGGGCTGAGCGTCCACGGAGCACGCTCTCG 2284			
by	181 GGAACATCATCGGCCGACGTGAAGGCACCTGATCCGGAGCTTCCCTGCCGATCGGGTCGGT 240			RESULT 6
Db	2285 GGCAACATCATCGGTCAAAGGGCCCCGGTGTGAGGGAGCAGGGCTCTGCCGCTCGGC 2344		ID ABN86383 standard; DNA; 3343 BP.	ABN86383;
by	241 TCACACTCGATTCTGTCGGAAACGGGGATGTTGATGGCACTGGCATCCATTGAG 300			XX ABN86383;
Db	2345 TCGCACTTCGACTCGGTCCGCAACGGGGATGTTGACGGCACCGCCTCG 2404			XX 21-OCT-2002 (first entry)
by	301 GCCCTGAGGCTGCCGGGTGATGGAAACGGCTACGTGAATGGCATCCATTGAG 360			XX Recombinant DNA encoding proteins involved in L-amino acid production.
Db	2405 GCGCTCGAGGCTGCGAGGGTGTGAGGGATATGTGAACCGTCATCCTCTCGAG 2464			XX DE
by	361 TTCATCGCGATCGTGGAGGAAGGGCCCCTCAGCAGTGGCATGTTGGGCCGG 420			XX KW Hydantoin racemase; HRase; bacterium; hydantoin; N-carbamoyl-amino acid; pharmaceutical; food additive; hydantoinase; HHase; Chase; gene; ds.
Db	2465 GTCAATCGCGATCGTGGAGGAAGGGGGCACCGCTTCAAGCGGGCATGCTGGGGTCGC 2524			XX OS Microbacterium liquefaciens.
by	421 GCCATTGCAGGGTTGGTCGCCGACAGGGAACTGGACTCTTGGTGTGATGAGGATGGAGTG 480			XX Key
Db	2525 GCGATCGGGGGCTGTCGCGAGGGCTGAGTCTGGACACCCCTGGTAGAACAGGGCTG 2584			XX CDS FT
by	481 TCCGTTAGGCAGGGGGCTACTGGCCTTGGCTTGAAAGGCCGGCAACTGCAGGGCTGCAGGCC 540			XX FT /tag= a
Db	2585 ACGGTGGGGCACGGGGCTTGGGAACGGGGTAGCTGGAGCGGGCTGCAGGGGGCC 2644			XX FT /product= "HRase"
by	541 CGCTCCGGGACCTGGCTTATCGAAACTACATGGAAACAAGGACCGATCCTC 600			XX FT /note= "hydantoin racemase"
Db	2645 CGTACGGGGATGACCTTCATCGAGTTGCACATCGAGGGGGCATGGGGGGCGATCCCTC 2704			XX FT 729. .2108
by	601 GAGCAGGAGCAAATAGAGATCGGAGTTGTAACCTCCATCGTTGGCTCATGGCG 660			XX CDS FT /tag= b
Db	2705 GAGCAGGAGAAAGGGAGATGGGGATCGGGGTCTGACGGGGATCGGTCTGGCTTCCGG 2764			XX FT /product= "HHase"
by	661 GTTGGCGTCAAAGGCCAACGGGACCCGGCAAAACCCCATGCACCTGGCCAGGAT 720			XX FT /note= "N-carbamoyl-amino acid-L-hydrolase"
Db	2765 ATCACGGTGAGGGCAGGGGGCTCATGGTGGAGGATCGGGGTCTGACGGGGCATGGG 2824			XX EP1188826-A2.
by	721 GCGCTGGTACCCGGCCTCTCATGGTGGAGGCTAACGGGTTCTGTCACAGGAGATCGCC 780			XX PD 20-MAR-2002.
Db	2825 GCGCTGGGGGGGGGGCTCATGGTGGAGGATCAATGGTTGGCTCAACGGAGATCGCG 2884			XX PP 09-AUG-2001; 2001EP-00119275.
by	781 GATGGCACAGTGGCTACCGTGGCCACCTCACAGTGGGGCTGGAGGCAACCAAGGTC 840			XX PR 13-SEP-2000; 2000JP-00278571.
Db	2885 GACGGCACGGGGGGGAGCTGAGGTGAGGGGGCTCAACCCCTGGGGCTCAACCGGTT 2944			XX PR 08-MAR-2001; 2001JP-00065815.
by	841 CGGGGGGGGGGACTTCACACTGGCTACCGTGGCCATGGGGCTGGCTGGCTGCC 900			XX PA (AJIN) AJINOMOTO CO INC.
Db	2945 CCCGGGGGGCTGGAGTCAACGGGGTGAACGGGGCTCAACCCCTGGGGCTCAACCGGTT 2944			XX DR WPI; 2002-510588/55.
by	901 CTGATCGACCGCATCTGGTCAACCTGGCTACCGTGGCCATGGGGCTGGCTGCC 960			XX DR P-PSDB; ABB80943, ABB80944, ABB80945.
Db	3005 CTGCTCGACAGGATCGGGGATGGGGATGGTGGCGAGTCGGCTCCGGGCC 3064			XX PT New 5-substituted hydantoin racemase, useful in production of optically active amino acids, comprises high working temperature, from Microbacterium species.
by	961 GATGTGGATGAATTTCAAATCTCAGCCCCGGCTGGCATGGTGGACGCC 1020			XX Disclosure; Page 30-31; 40pp; English.
Db	3065 GCGGTGAACGGGGTCTCTGGGCTCAGCGCTCTGGCTGGATGGATCGGC 3124			XX The invention relates to a 5-substituted hydantoin racemase (HRase) that is derived from a Microbacterium by culture, disruption or lysis, and purification. HRase has a high working temperature of 5-60 plusOC and is used to racemize optically active 5-substituted hydantoin for subsequent enzymatic conversion to N-carbamoyl-amino acids and then optically active amino acids (useful in pharmaceuticals, the chemical industry and as food additives). The present sequence represents a M. liquefaciens recombinant DNA containing structural genes encoding proteins such as HRase, HHase and Chase involved in L-amino acid production
by	1021 GTTCGCGAACGGGGCTCGGAACTCGGCTCACACCCGGGATATCAGCACTGGGGGG 1080			XX Sequence 3343 BP; 599 A; 1041 C; 1173 G; 530 T; 0 U; 0 Other;
Db	3125 GTGCGGACGGGGGGCTCGGAACTCGGCTCACCCATCGGACATCACAGGGGGAGGG 3184			XX 59.7%; Score 739.8; DB 6; Length 3343;
by	1081 CACGACTCGATGGTCACTGCCCAAGGTACGGACGTGGTAATGGTTTCCAAAGCCGT 1140			XX Best Local Similarity 74.8%; Pred. No. 7.7e-168;
Db	3185 CACGACTCGATGGTCACTGCCCAAGGTACGGACGTGGTTCCAAAGCCGT 1200			XX
by	1141 GCTGGCGGAGCCACGGTCCCGGAAAGAACATGGACCCCTTCGCAAAGGAACT 1200			Query Match

Matches	927	: Conservative	0	: Mismatches	312	: Indels	0	: Gaps	0
Qy	1	GTGACCCCTGCAGAAAGGCCAACGGCCATTGAGAAAGAGATCCGGGAGCTTCCCAGGCTCCTCCGG	60						
Db	2105	GTGACGGCTGCAGCGGCCGATGCCATCGAGGAGGCTCTGGACTCTCCCGC	2164						
Qy	61	TTCCTGGAGAAGGGCCCCGGTTACCCGGCTGAACCTCACACTCCAGAGCATGCCCGCG	120						
Db	2165	TTCCTGGTCAAGGGCCGGCTGACACGTCTCACGTACACTCCGGAGCACGCCGCG	2224						
Qy	121	CGGGAAACGCTCATGGGCTATGAAAGGGCCGGCTTGAGCGTTCTGTGAAGAGCCACTC	180						
Db	2225	CGAGAGGTGATCGTGGCCCATGGAGGGCTGAGGTCACAGGACGCTCTC	2284						
Qy	181	GGAAACATCATGGGGGACGTGAAGGGCACTGATCGGGAGCTCCTGCCATCGGGTCGGT	240						
Db	2285	GGAACATCATGGGCGGGTGAAGGGAGGGACCCGGCTTGGGATGCCGCTCGGC	2344						
Qy	241	TACACATTGATTCTGTCGGAAACGGGGGATGTTGATGGCACTGAGGGTGGTGC	300						
Db	2345	TCGCACTTGACTCGGTCGGCAACGGGGATGTTGATGGCATCCATTGAG	2404						
Qy	301	GCCCTTGAGGCTGCCCCGGTGTGGAAACGGCTACGTGAATGGCATCCATTGAG	360						
Db	2405	GCGCTGAGGCTGAGGGTGTGGAGAGGGATATGTGAACCGTCACTCTCGAG	2464						
Qy	361	TTCATCGCGATCGTGGAGGGAAAGGGCCGCTCAGCAGTGGCATCTTGGGGCG	420						
Db	2465	GTCATCGGATCGTGAAGGGCACCCGGTCAAGCAGGGATCTGGGGCATCTGGGTGCG	2524						
Qy	421	GCCATTAGGGTGGTGGAGGGAAAGGGGAACTGGGACTCTTGGTTGATGAGGAGGTG	480						
Db	2525	GGATCGGGCTGTTGGAGGGCACCCGGATCTGGACACCTGGGACAGGGGTG	2584						
Qy	481	TCCGTTAGGGAGGGCTACTGGCGGACAGGGAAACTGGGACTCTTGGTTGATGAGGAGGTG	540						
Db	2585	ACGGTGGGGATGACCCCTGGTCCGACGGCCACGGGGTGGAGGGGGGTG	2644						
Qy	541	CGCTCCGGGGGACCTGGCTGGCTTTATGAACTACACATGAAACAAGGACCGATCCCTC	600						
Db	2645	CGTACGAGGGATGACCCCTGGCTTCAATGAGTGGCACATCGAGTGGCACGGGGATCCTC	2704						
Qy	601	GAGCAGGAGCAAATAAGAGATGGGAGTTGTAACCTCCATCGTGGCATGGCG	660						
Db	2705	GAGCAGGAGAAAGTGGAGATGGGGATCGTGGTGTGGGGATCGTGGCTCCGG	2764						
Qy	661	GTTGCCGTCAAAGGAGAACGGGACGGGGACAACCCCCATGCACCTGGCAGGAT	720						
Db	2765	ATCACGGTGGGGAGGGGACCCGGGACGGGACCTGGACCCATGACGGGAGGAC	2824						
Qy	721	GGGCTGGTACCCGGCGCTCTATGGTGAGGGAGGTCAACCGGTCAACGGAGATCGCC	780						
Db	2825	GGCTGGTGGGGGCTCATGGTGAGAGATCAATGGGTCAACGGAGATCGCC	2884						
Qy	781	GATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGGGAGGTCAACGGAGTC	840						
Db	2885	GACGGCACGGGGGACCCCTCACGGTGGCATGGGGCTCAACGGGTT	2944						
Qy	841	CGGGGGAGGTGGACCTTCACACTGGACCTGGGTTCTCCGGCATGGAGTCGGT	900						
Db	2945	CGGGGGGGGGGAGGTGGACCTTCACACTGGACCTGGGTTCTCCGGCATGGAGTCGGT	3004						
Qy	901	CTGATCGACCGCATCTCGGTCAATGGTGGGGAGGTGGCTGGCTGGCG	960						
Db	3005	CTGGTCGACAGGATCGAGGGCATGGGCACTGGCTGGCTGGGGGGGGGG	3064						
Qy	961	GATGTGGATGAATTTCGAATCTCGGTCAATGGTGGGGAGGTGGCTGGCG	1020						
Db	3065	GGGGTGAACGGGTTCTGGGCTGAGGGCATGGGCTGGCTGGATCGC	3124						
Qy	1021	GTTCGCGAACGGGGCTCGGCCATTGGCATGGTGAATGGCAATGGC	1080						
Db	3125	GTGGCGAACGGGGTCCCGAACATCGGCTTCAACGGGAACTACGGGATG	3184						
Qy	1081	CACGACTCGATGTTCATGCCAGGTACGGGACGGTGGAAATGGTTTCAGGCCGT	1140	Qy					
Db	3185	CACGACTCGATGTTCATGCCAGATACCGACGGCTCCAGGCTCCTCCGC	3244	Db					
Qy	1141	GCTGGCCGGAGGCCACGGTCCCGAAGAATGGGACCTTCGATGACCTCGCAAAGGAAC	1200	Qy					
Db	3245	GGGGGGAGCCATGGCCGGAGGATGGTCCGATGGGGAGGATGGGGAGGAC	3304	Db					
Qy	1201	GAGGTTGTCCTCCGGTAATGAAGGCCATTGACCGGTAA	1239	Qy					
Db	3305	GATGGTGGCTTCAACGGTGGCTGACGGGGTTGACCGGTGA	3343	Db					

QY	373	GTTGAGGAGAAGGGCCCTTCAGCAGTGTGGCCATGTCAGGAGGG 432	WP	ABQ69245_06	600001	710000
Db	75532	GTGGAAGAAAGGCCTCTCGTTTGGCAGGAATTCAAGGCC 75591	WP	ABQ69245_07	700001	810000
QY	433	TGGTGGCCGACAGGGAACTGGACTCTTGGTGTAGGATGGAGTAGGCAG 492	WP	ABQ69245_08	800001	910000
Db	75592	AAAGTCACAACGAAATGCTAACAGAAATGAAAGATAAGATGGCATTACGGAT 75651	WP	ABQ69245_09	900001	1010000
QY	493	GCGGCTTAAGTGCCTTGAGGACTACAGTCAGGCTGCCGCTCAGGAGGA 552	WP	ABQ69245_10	1000001	1110000
Db	75652	GCCATGGCAAACCTAGGATTAAACGAAACGAGTACATAGGCCATTGGCACAAAGAG 75711	WP	ABQ69245_11	1100001	1210000
QY	553	GACCTGGTGTGCTTTATCGAACTACACATTGAAACAAAGGACCGATCCTCGAGCAGGAGCA 612	WP	ABQ69245_12	1200001	1310000
Db	75712	TCCGTGAAAGCATTATCGAACACTGCATATCGAACAGGCCCATCGAAAATGCCAGT 75771	WP	ABQ69245_13	1300001	1410000
QY	613	ATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGTGTGGGTGCGCATGTCAGGAA 672	WP	ABQ69245_14	1400001	1510000
Db	75772	GAAGACGTAGGGCTAGTCGATACAGTAGTCGGTTAACAAATCAACGTAACGTA 75831	WP	ABQ69245_15	1500001	1610000
QY	673	GGCAGAAGGACCCACGGGACAACCCCCATGCAACCTGCGCAAGGATGGCTGGTACCC 732	WP	ABQ69245_16	1600001	1710000
Db	75832	GGGCAAGGGCCATGCGAGGACAAACCCGATGCTTGACCGAAAGATGGCTTAGTTCC 75891	WP	ABQ69245_17	1700001	1810000
QY	733	GCGGCTCTCATGGTGAAGGGAGGTCAACCGGTTCTGCAACGAGATCGCCGATGGCACAGTG 792	WP	ABQ69245_18	1800001	1910000
Db	75892	GCTGTCGAATTAGGGCAATTGGCCAAATTGGCAGAACTCGTATCCAAAGAAGGGTAACTGTG 75951	WP	ABQ69245_19	1900001	2010000
QY	793	GCTACCGTTGCCACCTCACAGTGGCCCCGGTGGAGGCAACCAGGTCCCCGGGAGGTG 852	WP	ABQ69245_20	2000001	2110000
Db	75952	CTAACAGTCGGAAACTAAACGTCATCAATGGGCAAATGTTATTCCAGATAAAGTC 76011	WP	ABQ69245_21	2100001	2210000
QY	853	GACTTACACACTGGACCTCTGGCTTCCTCCGGCATGAGGAGCTGCTGCTGATCGACGCC 912	WP	ABQ69245_22	2200001	2310000
Db	76012	ATTTTACCGTGGACATTG --- CGCAAAAGACGAATTCACTGTTCAAAATACATTAGCA 76068	WP	ABQ69245_23	2300001	2410000
QY	913	ATCTCGGTCATGGTGGAGGTGCGCTCCAGGCCGGTGTGGCTGCCGATGTGGATGAA 972	WP	ABQ69245_24	2400001	2510000
Db	76069	AAAACGAAAGAAATCATTCAAGGCCACCGAAAAAAATGGTATTACTTGCGAAATAAAAGAT 76128	WP	ABQ69245_25	2500001	2610000
QY	973	TTTTTCAATCTAGCCGGTTCAGGTGGCTCCTACCATGGTGGACGGACTCGATG 1032	WP	ABQ69245_26	2600001	2710000
Db	76129	ATGATTACCAAAACCAACCCATTATCAAGCTTAAGTATCGGAAATAACGGCTAAGT 76188	WP	ABQ69245_27	2700001	2810000
QY	1033	GCCTCGCCCAAGGTCAACACCGGATATCAGCAGTGGGGCCACGGACTCGATG 1092	WP	ABQ69245_28	2800001	2910000
Db	76189	GCCGACCAATTGGCTTTAAGTATCGAACATAATGGTTAGCGGGGACCGCTAAGT 76248	WP	ABQ69245_29	2900001	3010000
QY	1093	TTCATCGGCCCAAGGTCAACACCGGATATCAGCAGTGGGGCCACGGACTCGATG 1152	WP	ABQ69245_30	3000001	3011208
Db	76249	ATTTTCGCTAGTTAACGAAAGTGGCTCTAGGTGATGGGGCTGATCTGAAATTCAGCA 76308	WP			
QY	1153	CACGTTCCGAAAGGAATGGACCGATTTCGATGACCTTCGCAAAAGGAACCTGAGGTGGCTC 1212	WP			
Db	76309	CATGCTCCAGAAGGAATGGACCGATTACGACAAGCTACAAAAGGAATTGAAGTCGTA 76368	WP			
QY	1213	CGGGTAATGAA 1223	WP			
Db	76369	AAGACAGTAAA 76379	WP			
QY	493	GCGGCTACTGCCTCGGCTTCAAGGAGCCGATCACATTGAAACAGGAGGG 552	WP			
Db	69277	GCTATGGCAAACCTAGGATTGACGCCATTCAAGTTCACTGGTTAAAGAT 69276	WP			
QY	553	GACCTGCGTGCCTTATCGAACTACACATTGAAACAGGAGGG 612	WP			
Db	69337	TCGGTCAAAAGCTTCATTGAACTACATGCCAACAGGCCAGTCTGCGCTGGGTCAA 69396	WP			
QY	613	ATAGAGATGGAGTTGCTCTCGGCTTCAGTGGCTGGCATTTGGGGTTGGGTCAA 672	WP			
Db	69397	GAAGATGTTGGCTAGTTGACACGGTCTGGTTAACAGAAATAAAAGTTACCGTAAAG 69456	WP			
QY	673	GGCAGAAGGACCAAGGCCAACCCCCATGCCAGGATGGCTGGTACCC 732	WP			
Db	69457	GGTCAAGGAGGTCAAGCCAAACTACGCCAACAGAAATGGCTCTAAAGTGCA 69516	WP			

RESULT 8
ABQ69245_05
Continuation (6 of 31) of ABQ69245 from base 500001 (Listeria innocua DNA sequence #684.
WP Sequence split into 31 fragments LOCUS ABQ69245 Accession Abq69245
WP Fragment Name Begin End

ABQ69245_00 1 110000 210000 310000 410000 510000 610000
ABQ69245_01 100001 200001 300001 400001 500001
ABQ69245_02 200000 310000 410000 510000 610000
ABQ69245_03 300001 410000 510000 610000
ABQ69245_04 400001 510000 610000
ABQ69245_05 500001

QY	733	GCGCTCATGGTGAAGGTCAACCGGATGCCGATGGCACAGTG	792	Db	78831	GCTAGCGTTTCATGAACAGCAATAAACCCATTTCACCTGAATCATGCCAATG	78890
Db	69517	GCTGCCAAATTCTTAATAAGCTTCCCTGAACTAGCTATCCAGAACGGAAACTGTC	69576	QY	373	GTGGAGGAGGAAGGGGCCCTTCAGCAGTGGCATGTTGGCGGCCATTGCCAGGG	432
QY	793	GCTACCGTGGCACCTCACAGTGGCCCCGGGAGGGTACCGGAAAC	852	Db	78891	GTGGAAGAAGGGATCTCGTGTGGAGGGACTACTTGCCTCCGAACAAATTACAGGA	78950
Db	69577	TTAACAAATTGGCAAACTTAATGTCTACCCGAACGGCTAAATGTAAT	69636	QY	433	TGGTGGCGACAGGGAAACTGGACTCTTGGTGTGAGGATGGAGTGGTAGGAG	492
QY	853	GACTTACACTGGACCTGGCTTCTCCGCGATGAGGAGTCGACGCC	912	Db	78951	AAAGTTACAAAAGAAATGGTACATGAAATGAAAGATAATAATGGGTTACTGCCAGAA	79010
Db	69637	GTTCACAGTAGATATTCAAGGCAAAATTCCACGTCCAAAATAGAAAAA	69696	QY	493	GCGGCTACTGCGCTTGGCTTGAAGGCCGGGAAACTGCAAGGTGCAAGGGTCCGGGG	552
QY	913	ATCTCGGTCACTGGTCGGCGAGGTGCGCTCCAGGCCGATGTGGATGAA	972	Db	79011	GCTATGGCAAACCTAGGATTGACGAAATCAAGTCATAAGCCATTGGTCAAAGAT	79070
Db	69697	A---CAAAAAAGTTATTCAAGCTTTGAAAGGGTATTATGTGTGAAATAAGAGAT	69753	QY	553	GACCTGGCTGCTTTATGAAACATACATTGAAACAAGGACCGATCTCGAGCAGGAGCA	612
QY	973	TTTTCAATCTAGCCCCGGTGCAGTGGCTTACATGGTGGACGCCGTGCGGAAGGG	1032	Db	79071	TCCGTCAAAGCTTCTATTGAACATCACATGAAACAAGGCCAGTCTAGAAATGCGAAAT	79130
Db	69754	ATGCTCTATGAAAACCAACATTTAACAGAAATTCAAGCAATTGACCGAAAGC	69813	QY	613	ATAGAGATGGAGTTGTAACCTCCATGTTGGCTTCGGCATTGGGGTGTCCAAA	672
QY	1033	GCCTCGGCCCTTGCAGTTCACACACGGGATATCAGCAGTGGGGGGGCCACGACTCGATG	1092	Db	79131	GAAGATGTTGGCTAGTTGACACGGTCAGCCATTGCTAGAAACGAAATAAGTAAAG	79190
Db	69814	GCGATAAAACTTGGCTTAAATATCGAACAAATGGTTAGCGGGCCAGGACACGATGCAATG	69873	QY	673	GGCAGAAGGACCAACCCCATGACCTGCGCCAGGATGGCTTACAGTGACAGTG	732
QY	1093	TTCATGCCCAAGGTCACGGAGCTGGCTTCAAGGGGATTCAGCAGTGGGGGGGGAGC	1152	Db	79191	GGTCAAGGAGGTCAAGCCGGAAACTACCCCATGCTAGCTAAAGTGTCAAGTGC	79250
Db	69874	ATTTTCGCAAGTTAACCGAAGTAGGACTTATTGTACCAAGCCACAAAGGTATAAGC	69933	QY	733	GCGCTCTCATGGTAGGGAGGTCAACGGGTTCTGTCAACGAGATGGCTACAGTG	792
QY	1153	CACGTTCCCGAAGGAATGGACCGATTTCGATGACCTTCGCAAAGGAACACTGAGGTTGTCT	1211	Db	79251	GCTGTCCAAATTCTTAAATAGCTTCTGAACTAGCTATCCAGAAGGGGGAACTGTC	79310
Db	69934	CATGCCACCAAGAAGGAATGGACCGATTACGACAAGCTCCAAAAGGCATCGAAGTCGTA	69992	QY	793	GCTACCGTTGGCCACCTCACAGTGGCCCCGGTGGAGGCAACCAGGTCCCGGGGAGGT	852
Db	ABQ67197_04	TTAACAAATTGGCAAACCTTAATGTCTACCCGAACGGCTAATGTAATAACAAACAAAGTC	79311	Db	79311	TTAACAAATTGGCAAACCTTAATGTCTACCCGAACGGCTAATGTAATAACAAACAAAGTC	79370
QY	853	GACTTCACACTGGACCTGCCTTCCTGGCATGAGGAGTGGCTCCGGGTGCTGACGCC	912	QY	853	GACTTCACACTGGACCTGCCTTCCTGGCATGAGGAGTGGCTCCGGGTGCTGACGCC	912
WP	Sequence split into 12 fragments LOCUS ABQ67197 Accession Abq67197	End	Db	79371	GTTTCACAGTAGATATTCAAGGAAATACTCCACGTCAAATACTAGAAAAA	79430	
WP	Fragment Name	Begin	QY	913	ATCTGGTCATGGTGGCGAGGTGGCTCCAGGGCGGTGGCTGGATGAA	972	
WP	ABQ67197_00	1	Db	79431	A---CAAAAAAGTTATTCAAGGTTTCGAAAGGAAATACTAGAAAGAT	79487	
WP	ABQ67197_01	100001	QY	973	TTTTCAATCTAGCCGGTGGCTCCTACCATGGTGGACGCCGTTGGCTGGCAAGGC	1032	
WP	ABQ67197_02	200001	Db	79488	ATGCTCTATGAAAAACACATTATCAACAGAAATTCAAGGATTTGTAAGGAAAGC	79547	
WP	ABQ67197_03	300001	QY	1033	GCCTGGCCTTGCAGTTCACACCCGGATATCAGGAGTGGGGGGGCCACGACTCGATG	1092	
WP	ABQ67197_04	400001	Db	79548	GCGATAAAACTTGGTTAAATATGAAACATGGTTAGGGCAAGGACACGATGCAATG	79607	
WP	ABQ67197_05	500001	QY	1093	TTCATGCCACGGTCAGGTTCACACCCGGATATCAGGAGTGGGGGGGCCACGACTCGATG	1152	
WP	ABQ67197_06	600001	Db	79608	ATTTTCGCAAGTTAACCGAAGTAGGACTTATTGTTGTAACAGGATATAAGC	79667	
WP	ABQ67197_07	700001	QY	1153	CACGTTCCCGAAGGAATGGACCGATTGGCTTCCGAAGGAACCTGGTTGCTCCT	1211	
WP	ABQ67197_08	800000	Db	79668	CATGCACCAAGAAGGATGGACCGATTAGCAAGGCTCCAAAAGGCATCGAAGTCGTA	79726	
WP	ABQ67197_09	900000	QY	RESULT 9			
WP	ABQ67197_10	1010000	Db	AAQ26728			
WP	ABQ67197_11	1100000	QY	AAQ26728			
WP		1110000	Db	AAQ26728			
WP		1163000	QY	AAQ26728			
Query Match	17.8+	Score 220.6; DB 6; Length 110000;	Db	AAQ26728			
Best Local Similarity	50.2%	Pred. No. 1.1e-42; Mismatches 564; Indels 3; Gaps 1;	Db	AAQ26728			
Matches 572; Conservative 0; MisMatches 564; Indels 3; Gaps 1;			QY	1153	CACGTTCCCGAAGGAATGGACCGATTGGCTTCCGAAGGAACCTGGTTGCTCCT	1211	
QY	73	GGCCCCGGTGTACCCGGTACACTCCAGAGCATGCCGCGGGAAACGCTC	132	Db	79668	CATGCACCAAGAAGGATGGACCGATTAGCAAGGCTCCAAAAGGCATCGAAGTCGTA	79726
Db	78591	GGTCAGGGAAACAACTGGCTTACATAGGAAAGAAGACCTGGGGCTTAATTATTAA	78650	QY	133	ATTGGGGCTATGAAAGGCCGCGCTTGAAGGAGGCACTCGGAAACATCATC	192
Db	78651	AAAGAAGAAATGGCTAAAGTAGGGCTTACTGTCTCAGAAAGATGGGATTGAAATATC	78710	Db	78651	AAAGAAGAAATGGCTAAAGTAGGGCTTACTGTCTCAGAAAGATGGGATTGAAATATC	78710
QY	193	GGCGGAGCTGAAGGCACCTGATCCGGAGCTTCCCTGGATTCACACTTGGAT	252	QY	193	TCTGTCGAAACGGGGATGTTGATGGCACTGGCAGGGTGTGGCTGGCT	312
Db	78711	GGACGACTAGAAGGGGACAATCCAGACATACAGGAGTAATGGCTGGTTCTCATTTGAC	78770	Db	78711	TCTGTCGCAAAATGTTGAGGTCCAGGAGTAACTGGCTGAAAGTAA	78830
QY	313	GCCCCGGTGTGCTGGAGAACGGCTACGGTAAATGGCATTCATTGAGTTCATCGCGATC	372	QY	313	GCCCCGGTGTGCTGGAGAACGGCTACGGTAAATGGCATTCATTGAGTTCATCGCGATC	372

Heat resistant carbamylase gene.
Temp; stable; N-carbamyl-Lamino acid; ss.
Escherichia coli.

XX	Key	Location/Qualifiers	
FH	CDS	320. .1550	
FT		/tag= a	
FT			
XX	JP04183391-A.		
PN			
XX	PD	30-JUN-1992.	
XX	PF	15-NOV-1990;	90JP-00307221.
XX	PR	15-NOV-1990;	90JP-00307221.
XX	PA	(NIPS) NIPPON SODA CO.	
XX	PT	Heat resistant carbamylase - is used to produce L-aminoacid in high purity by reacting enzyme protein with N-carbamyl-L-aminoacid.	
XX	DR	WPI; 1992-265588/32.	
XX	DR	P-PSDB; AAR25693.	
PS			
XX	CC	The DNA encodes a heat resistant carbamylase (optimum temp. ca. 60 degrees C, optimum pH ca. 8.5) which can be stably produced by bacteria.	
CC	CC	The enzyme allows efficient prodn. of a high concn. of N- carbamyl-L- amino acids which may be hydrolysed to the relavent L- amino acids in the presence of the heat resistant carbamylase	
CC	CC		
SQ	Sequence	1860 BP; 438 A; 456 C; 580 G; 386 T; 0 U; 0 Other;	
XX	CC	Score 182.2; DB 2; Length 1860;	
CC	CC	Pred. No. 7.1e-34;	
Matches	564; Conservative	0; Mismatches 593; Indels 6; Gaps 2;	
QY	72	AGGCCCGGGCTTACCCGGCTGACCTTACACTCCAGAACATGCCGCATGGCTGAAAGCCT 131	
DB	382	AAGGGGGCGCTCACGGCCTCTCGTTCACTGCTGAAGAGCGGGCCAAGATCTGT 441	
QY	132	CATTGGGCTATGAAAAGGGCGCCRTGAGCGTTICGTGAATGAGAACATCAT 191	
DB	442	CGCTTCCTACATGGCGGAAGGCCGGCTTTCGTATATGAAGAACGGCTGGCAACTTGAT 501	
QY	192	CGGCCGACGTGAAGGCACTGATCCGGAGCTTCCCTGGGATCGGGATCAGCTTCGA 251	
DB	502	CGGACGAAAGGAGGGAATCCGGATGCCACGGTGTCTTGTGGATCTCATCTCGA 561	
QY	252	TCTGTCCGAAAGGGGATGTTGATGGCACTGAGGCTACGTGAATCGGCATCCATTGAGGC 311	
DB	562	TTCGGTTACAACGGGGCTGCTTGTGGACCGTGGGGTGAATTGAAGTAGTGGGTT 621	
QY	312	TGCCCGGTGATGGCTACGTGAATCGGCATCCATTGAGGTCACTGGGAT 371	
DB	622	CGTTCAGACGATGAACGGGACCGGACCCATTGGGGTGGGGTGAATTGAAGTAGTGGGTT 681	
QY	372	CGTGGAGGAGGAAGGGGGCGCTTCAAGCACTGGCATGTTGGGGCATTGCAGG 431	
DB	682	CACTGAGGAAGGGAGGGGGCGCTTGTGGCATGATGGCAAGGCCCATGGGG 741	
QY	432	GTTGGTGGCGACAGGAACACTGGACTCTTGGTGTAGGAGTGGAGTGGCTTAGGCA 491	
DB	742	AAACACTGCC--GCCGGAAGGCCGCTGGAGTGGGGGATTCCCTCGCTGA 798	
QY	492	GGACCTGGTGGCTTTATCGAAACACTACACATTGAAAGCAAGGACCGACCA 611	
DB	799	AGCGATGAAAAGGGCTTGAACCCGGACCGCTTGCCTGGAGGACAGG 858	
QY	552	GGACCTGGTGGCTTTATCGAAACACTACACATTGAAAGCAAGGACCGACCA 611	
DB	859	AACGGTGAAGGCCATATGTCGAATTGCAATCGAAACAGGACGGTGGCTGA 918	
QY	612	AATAGAGATCGGAGTTGTAACCTCCATCGTTGGCGCTTGCCTGGCATATGCA 671	

XX	(RHON) RHONE POULENC NUTRITION ANIMALE.	QY	732 CGCCGCTCATGGTAGGGAGTCACCGGTTGTCAACGAGATGCCGATGGCACAGT 791
PA		Db	720 GGCGCCAGATCATAGTGTACGGAAAGGAAAGCAGGAAACGGCAACGGT 776
XX	Dion M, Batisse N, Weigel P, Lecocq FM, Hallet JN, Sakanyan V;	QY	792 GGCTACCGTGGCCACCTCACAGTGGCCCGGCAACCAAGGTCCCGGGAGGT 851
PI		Db	777 CGGTACTGAGTAGGCATGTATAATGTCATTCCGGTAACCGGT 836
XX	DR; WPI: 1996-335881/34.	QY	852 GGACTTCACACTGGACCTCGCATGAGGAGTCGCTCCGGTGTGATGACCGG 911
DR	P-PSDB; AAW03544.	Db	837 CGAAATTGGCTCGATTGGCAGCTTGAAAGGTGATGAAAGC 896
XX	New stereospecific, heat-stable amino acid amidohydrolase - from B. stearothermophilus and related DNA, esp. for prodn. of L-Met.	QY	912 CATCTGGTCATGGTGGAGGTCGCTCCAGGGCGGTGGATGAA 971
PS	Claim 4; Page 9-11; 17pp; French.	Db	887 CATAGCCGTCGGGAGAGACATGCCAAGGAGCGAACGTTGGCTCACGACCGAGCG 956
XX	PstI fragments of <i>Bacillus stearothermophilus</i> NCIB 8224 genomic DNA were inserted into pBR322 and used for transforming E.coli hosts. Selection was on medium which contained tetracycline but lacked arginine. All positive transformants carried plasmids with a 4.7 kb insert. The present sequence is a fragment of the insert. Amino acid amidohydrolase encoded by the fragment is stereospecific and heat-stable (maximum activity at 55 °C -60 deg.C). The enzyme hydrolyses N-carbamoyl amino acid derivs. to L-amino acids and is particularly useful for production of L-methionine. (Updated on 16-OCT-2003 to standardise OS field)	QY	972 ATTTTCAAATCTCAGGCCGGTGCAGCTGGCTACCATGGTGGACGCCGTTGCGAAGC 1031
CC		Db	957 ACTGCAAGAAATGGCCGGTGTATGTTCGAGGTGGTAAACAGGCCGGAAGAGC 1016
CC		QY	1032 GGCTCTGGCTTGCAGTTCACACACCGGGATATCAGCAGTGGGGGCCACGACTCGAT 1091
CC		Db	1017 GTGCAAGGAGCTCGGGTATCCGGGTTGCTGCCAGGGCGCAGCCCCATGACGGGT 1076
XX	Sequence 1230 BP; 268 A; 312 C; 413 G; 237 T; 0 U; 0 Other;	QY	1092 GTTCATGCCAGGTACGGACCGTGGAAATGGTTTCAGCTGGCTGGCCGGAG 1151
Query	Match Score 179; DB 2; Length 1230;	Db	1077 ACAGTTGGCTCCGATTTGGATGATGGATTTGTCGCTCCCAGGGT 1136
Best Local Similarity 48.3%; Pred. No. 3.8e-33;	Mismatches 595; Indels 6; Gaps 2;	QY	1152 CCACGTTCCCGAAGAATGGACCCATTGATGACCTTGCACAGGAAACTGAGGTGTCCT 1211
Matches 562; Conservative 0; Mismatches 595;	Db	1137 TCATAGTCGGGGATGGAGTACTAAAGAAGACTGCGCCGTTGGAGGTA 1196	
QY	72 AGGCCCGGTTACCCGGTGTGACCTACACTCCAGAGCATGCCGGCGGGAAAAGCT 131	QY	1212 CCGGTAATGAAGGGCACTTGACCC 1234
Db	63 AAGGGGGGTCAAGGGGCTCTCGTTCACTGCTGAAGAGGGGCAAAAGATCTCGT 122	Db	1197 TCATACAGTGTGGCAACTGGCC 1219
QY	132 CATTGGGGCTATGAAAGGGGGCTTGTAGGCGTTCTGTAAGACGGCAACTCATCAT 191	RESULT 12	
Db	123 CGTTTCTACATGCGCAAGGGGGCTTTTCGTATATGAAAGACGGGGTGTGCAACTTGTAT 182	ID	ABQ71004/c
QY	192 CGGGCGACGTGAAGGCACTGATCGGAGCTTCCGTGATGGTTCACACTTCGA 251	Db	ABQ71004 standard; DNA; 3880 BP.
Db	183 CGGACGGAAAAGGGACGAATCGGATGCCAGGATCGGATTCATCTCGA 242	XX	XX
QY	252 TTCTGTCCGAAACGGGGATGTTGATGGCACTGCAAGGGCTTGTGCGCCCTTGAGGGC 311	AC	ABQ71004;
Db	243 TTGGTTTACAACGGGGCTGCTTGTGGACCGCTGGTGGATCTCATCTCGA 302	XX	XX
QY	312 TGCCGGGTGATGCTGGAGAACGGCTACGTGAATGGCATTCATTGAGTTCATCGGAT 371	DT	29-AUG-2003 (revised)
Db	303 CGTTCAGACGATGAACGGACACGGAACCCZATTGAAGTAGTGGCGGT 362	DT	29-AUG-2002 (first entry)
QY	372 CGTGAGGAGGAAGGGCCCGCTTCAAGTGGCATGTTGGCTGGCCGGCATTGCGAG 431	XX	XX
Db	363 CACTGACGAAGGGAGGGCTTCGTGTCGGCATGATGGCAGCCGCGCATGGCG 422	DE	Listeria monocytogenes 4b contig DNA sequence #946.
QY	432 GTGGCTCGCCGACAGGGAACTGGACTCTTGGTGTAGGATGGAGTGGCTTAGGCA 491	XX	XX
Db	423 AACACTGCC--GCCGGAAAGGGCTGAGGTGGCCGAGCGGGAAAGGGATTCCCTCGTGA 479	OS	Listeria monocytogenes ATCC 19115.
QY	492 GGCGGCTACTGCCCTGGCTTATCGAACTACACATTGAACAGGGCTCCGGGGC 551	XX	KW
Db	480 AGCGATGAAACGGGGCTGTGACCGGGCTGGCTGGAGGAAACCCAGGG 539	XX	infection; ds.
QY	552 GGACCTGGCTTATCGAACTACACATTGAACAGGGCTCTCGAGGCCGAGGG 611	XX	XX
Db	540 AACGGTGAAGGCTATGTCGATTGCAATATGAAACAGGGCTGGAGGGCTGG 599	PR	WO20028891-A2.
QY	612 AATAGAGATGGAGTTGTAACCTCCATGTTGGCTGGCATGGCTGGCAA 671	XX	04-OCT-2001; 2001WO-FR003061.
Db	600 TCTTCCAGTGGCATCGTCAGTGGCATGCCGGTGAATTTACCATCGC 659	XX	04-OCT-2000; 2000FR-00012697.
QY	672 AGGCAGGAAGGACACGCCGGCACACCCCATGCACTGCGCCAGGATGCGCTGG 731	PA	(INSP) INST PASTEUR.
Db	660 CGGCGGGGGAAATGCCGACGCCGATGTCATTGGCGACCGATGGGGC 719	PA	(CNRS) CNRS CENT NAT RECH SCI.
QY		PI	Kunst F, Glaser P;
Db		XX	WPI; 2002-332479/37.
		DR	
		XX	New genomic sequences from <i>Listeria</i> species, useful for detection, treatment and prevention of infection, also related polypeptides, antibodies and modulators.
		PT	
		PT	
		PS	Claim 14; SEQ ID NO 3817; 180pp; French.

The present invention relates to nucleic acid sequences (ABQ67188-ABQ71212) from Listeria sp. The sequences are useful as probes and primers for identification and/or detection of Listeria (e.g. as contaminants in foods, or mutational analysis) and for analysis of gene expression. Proteins encoded by the nucleic acid sequences can be used to screen for compounds that modulate gene expression, replication and pathogenicity of Listeria (potential therapeutic agents), also for treating infections by Listeria, and are useful as immunogens in anti-Listeria vaccines. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences). (Updated on 29-AUG-2003 to standardise OS field)

XX Sequence 3880 BP; 988 A; 777 C; 810 G; 1304 T; 0 U; 1 Other;

Query Match Score 12.8%; DB 6; Length 3880;
Best Local Similarity 49.8%; Pred. No. 4-e-28;
Matches 429; Conservative 0; Mismatches 429; Indels 3; Gaps 1;

Qy 351 TCCATTGAGTTCATGGGATCTGTGGAGGAAGGGCCGTTCAAGCATGGT 410
Db 3855 TCGCTTGAAATTATTGGGATGGTTGGAAAGAAGAAGGGCCTCGTTGGCTCCT 3796
Qy 411 GGGGGCGGGCATTGCAAGGTGGCTGGCGACAGGGAAACTGGACTCTTGGTGATGA 470
Db 3795 TGCTTCACGGTACATTACTGCCAAAGTTACCCAAGAAATGCTCCATGAAATGAAAGATA 3736
Qy 471 GGATGGAGTGTCCGTAGGCAGGGCTACTGCCCTGGCTTGAAACTACACATTAACAGGG 530
Db 3735 AGATGGAAATTACAGCCGCAAGGCCATGGCAAAGTAGGATTGATGCAAATCAAGTAGT 3676
Qy 531 GGCTGCAGCCCCGCTCTGGCAGGGGACCTGGTGTCTTATCGAAACTACACATTAACAGGG 590
Db 3675 TACAGCTATCCGACGAAAGAATCCGTTAAAGCATTATCGAAACTACACATTAACAGGG 3616
Qy 591 ACCGATCCTCGAGCAGGGCAAATAGAGATCGGAGTTGTAACCTCCATCGTTGGGTTTG 650
Db 3615 ACCAGTCCTAGAAAATGCCAAATGAAAGATGGGCTAGTACAGTAGTTAAC 3556
Qy 651 CGCATGGGGTTGCCGTCAAAGGCAGAAGGCCAACGCCAGCGGGCAACACCCCCTGACCT 710
Db 3555 AGAAATAAAAGTAACAGTAAAGGACAAGGGCAAGGCCATGCCAGGCAACACCCCCTGAA 3496
Qy 711 CGCCCGGGATGCCGTGGTACCGCCGCTCTCATGGTGAGGGAGGTCAACCGGGTCTGCAA 770
Db 3495 CGAAAAGATGCCGTGATCACAGCTGTCAAATTAGGTCAAACACTACCAGAACTAGCTAT 3436
Qy 771 CGAGATGCCGATGGCACAGTGGTACCGTTGGCCACCTCACAGTGGCCCCGGGG 830
Db 3435 CCAAGAAGGGGGAACTGTATTAAAGTGGCAAAACTCAACGTCTATCCAATGGGC 3376
Qy 831 CAACCAAGGTCGGGAGGTGGGACTTCACACTGGACCTCTCCGATGGAGTC 890
Db 3375 AACAGTTATTCCAGATAAGTGTGTTTACCGTAGATATCGGAAAGACGAAATTCA 3316
Qy 891 GCTCCGGCTGCTGATGACCGCATCTGGTCATGGTGGAGGTGGCTCCAGGGCG 950
Db 3315 CGTCCAAATAATAGCAAAAGAAATTAT--CCAATCCGAGAGAAAACGG 3259
Qy 951 TGTGGCTGCCGATGTGGATGAAATTTCATCTCAGCCGGTGCAGCTGGCTCTACCAT 1010
Db 3258 CATCACTTGCAGAAATAGAGATAATGCTATAAGAACCCGCGACCATTTAAAGAGAT 3199
Qy 1011 GGTGGACGGCCACGACTGATGTTGGAAGGGCCTCGCAGTTCACACACGGGATATCAGCAG 1070
Db 3198 TCATCAAGCATTGAAAGTGGCCACAACTCGGCTTAAATACCGAACAAATGGTTAG 3139
Qy 1071 TGGGGGGCCACGACTGATGTTGGAAGGGCCTCGCAGTTCACACACGGGATATCAGCAG 1130
Db 3138 CGGGGGGACAGATGCAATGATTGGTTAACCGAACGGTGGCTGATGGCTTGT 3079
Qy 1131 TCCAAGCCGTGCTGGCGGAGGCCACGTTCCGAAAGAATGGACCTTCGATGACCTTCG 1190

XX The present invention relates to nucleic acid sequences (ABQ67188-ABQ71212) from Listeria sp. The sequences are useful as probes and primers for identification and/or detection of Listeria (e.g. as contaminants in foods, or mutational analysis) and for analysis of gene expression. Proteins encoded by the nucleic acid sequences can be used to screen for compounds that modulate gene expression, replication and pathogenicity of Listeria (potential therapeutic agents), also for treating infections by Listeria, and are useful as immunogens in anti-Listeria vaccines. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences). (Updated on 29-AUG-2003 to standardise OS field)

RESULT 13
AAQ15312
ID AAQ15312 standard; DNA; 7475 BP.
XX XX
Db 3078 CCCTAGGCCACAATGGTATAAGCATGGTTGTCTT 1211
Qy 1191 CAAAGGAACCTGAGGTTGTCTT 1211
Db 3018 AAAAGGAATCGAAGTCGTACT 2998
XX AC
Qy XX
Db 27-AUG-2003 (revised)
Db 25-MAR-2003 (revised)
Db 17-MAR-1992 (First entry)
XX DE Plasmid PHPB12 encoding enzyme able to transform 5-substd. hydantoin into N-carbamyl amino acid.
XX KW Hydrolase; SS.
XX OS Pseudomonas sp.
XX FH Key
CDS FT CDS 1156. .3228
/*tag= a
/label= ORF 2
FT CDS 3232. .5010
/*tag= b
/label= ORF 3
FT CDS 5031. .6275
/*tag= c
/label= ORF 4
XX PN JPO3251176-A.
XX PD 08-NOV-1991.
XX PR 10-JAN-1990; 90JJP-00002786.
XX PA (NIPS) NIPPON SODA CO.
XX DR WPI; 1991-373414/51.
P-PSDB; AAR15482, AAR15483, AAR15484.
XX PT Enzyme for transforming 5-substd. hydantoin - in which enzymatic protein has specified aminoacid sequence.
XX PS Disclosure; Fig 3; 11pp; Japanese.
XX CC The sequence comprises three open reading frames: 2, 3, and 4, which together encode an enzyme which has the ability to transform 5-substd. hydantoin into N-carbamyl amino acids. The DNA segment contg. the three ORFs was obtnd. from Pseudomonas NS 671 and ligated into pUC18 to form the plasmid PHPB12. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 27-AUG-2003 to correct OS field.)
XX SQ Sequence 7475 BP; 2424 A; 1237 C; 1698 G; 2116 T; 0 U; 0 Other;
Query Match 12.8%; Score 158; DB 2; Length 7475;
Best Local Similarity 48.0%; Pred. No. 6.4e-28;
Matches 514; Conservative 0; Mismatches 550; Indels 6; Gaps 2;
Qy 142 ATGAAAGGGCCGCCTGAGCGTCTGGCTGAAGACGCCTGGACATCGGCCGACGT 201
Db 5175 ATGAGAGGGCCGCCTTACTGTRTACTCATGATCATTTGGAAATTGAGCGAAA 5234
Qy 202 GAAGGCCACTGATCCGGAGCTCCCTGCGATCGGGTCAACACTCGATTCTGRCGGA 261

Db	5235	GAGGGAGAAACCCCAAGCCTTCCTCCGTAATGATCGGTTCTCATATTGATTCTGTCGA	5294		DT	13-AUG-2002	(first entry)
Qy	262	AACGGGGATGTTGATGGCACTGCAGGGCTGGTGTGCCCTTGAGGTGCCCCGGGTG	321	XX	DE	Bacillus licheniformis	genomic sequence tag (GST) #308.
Db	5295	AATGGGTTAAGTTGATGGGTATTGGAGTACTGGCAGGGATTGAATTCGTCACGCC	5354	XX	KW	Differential gene expression; genomic sequenced tag; GST;	
Qy	322	ATGCTGGAAACGGCTACGGTAATGGCATCCATTGAGTTCATCGGATCGTGGAGGAG	381	KW	altered culture condition; environmental stress;		
Db	5355	ATATCCGAGGCAAATGTCATGAACATTCAATGAAAGTGGTAGTTTGTGAAGAG	5414	KW	physiological provocation; ds.		
Qy	382	GAAGGGCCGCTTCAGCAGTGGCATGTTGGGGCATTCATTGAGGGTTGGTGC	441	XX	OS	Bacillus licheniformis.	
Db	5415	GAAGGTTCAAGTTAAATGACGGATTATTGGAAGTAGGGCATGGTAAAGTAA	5474	XX	XX	PN	WO200229113-A2.
Qy	442	GACAGGAACCTGGACTCTTGGTGTAGGGATGAA-GTGTCCGTTAGGCAGGGGCT	498	XX	XX	PF	2001WO-US031437.
Db	5475	CCCGAGGATTACAAAAGTGTACGATAACAATGAACTGGTAAAGGCTAA	5534	XX	PR	06-OCT-2000;	
Qy	499	ACTGCCCTGGCTGAAGGCCGGGAACCTGCAAGGTGCTCGGGGGACTCTG	558	XX	PR	2000US-00680598.	
Db	5535	TGGCTTGGGATAGATCCGATTTACCCACCATCTACGGAGATGGTGACATT	5594	XX	XX	PR	27-MAR-2001;
Qy	559	CCTGCTTTATGAACTACACATTGAAACAGGACCGATCCTCGAGGAGCAAATAGAG	618	XX	PA	NOVOZYMES BIOTECH INC.	
Db	5595	AAACATTTATTTGAGATGCAATTGAGCAAGGGCATATCTAGAAAAGATAATTACCA	5654	XX	PA	(NOVO) NOVOZYMES AS.	
Qy	619	ATCGGAGTTGTAACCTCCATCGTTGGCTTCGGCATGGTCAAAGGGAGA	678	XX	XX	PA	Monitoring differential expression of several genes in first Bacillus
Db	5655	ATAGGGATTGTAAGTGGGATCTGGGTCTCTGGTTCAAGGTAAGGCTAGTTGGAGAA	5714	XX	PT	cell relative to expression of same genes in one or more second Bacillus	
Qy	679	AGCGACCAAGCCGACAAACCCCCATGCAACCTGCCAGGATGCCGTGGTACCCGCC	738	XX	PT	cells, by using substrate containing Bacillus genomic sequenced tag	
Db	5715	GCTGCCATGGGGACTGTCCCTATGAGTCTACGCAAAGACCCCCTAGTGGGGCTGCG	5774	XX	PT	array.	
Qy	739	CTCATGGTGAGGGAGT--CAACCGGTTCTGTCACAGGAGATGCCGATGGCACAGTG	795	XX	XX	Claim 4; SEQ ID NO 308; 200pp; English.	
Db	5775	GAAGTCATTAAGAAGTAGAAACTCTTGTATGAAAGTCCAAATGCTCCACAGTGGC	5834	CC	The invention describes a method of monitoring differential expression of		
Qy	796	ACCGTTGGCCACCTCACAGTGGCCCCGGTGGAGGCAACCAGGTCCGGGGAGGTG	855	CC	genes in a first Bacillus cell relative to expression of the genes in		
Db	5835	ACTGTTGGTAGAAATAGCGGCTTTCTCTGAGGAAGTAATATTATCCTGATCAGTGG	5894	CC	other Bacillus cells, comprising hybridising labelled nucleic acid probes		
Qy	856	TTCACACTGGACCTGCGTTCCGCATGAGGAGTCGCTCCGGCTGCTGATGACCG	915	CC	isolated from Bacillus cells to a substrate containing array of Bacillus		
Db	5895	TTTACCTGTGATATCAGGGATATTGAACTTGAAACTTGAAATTATGAGAGATA	5954	CC	genomic sequenced tags (GST), examining the array, and determining		
Qy	916	TGGTCAATGGTGGCAGGGTGGCTCCAGGGCGGTGGCTGCCATGGTGGATGAA	975	CC	relative gene expression by an observed hybridisation reporter signal of		
Db	5955	GAAGAAAAGATCAAACCTGTAAGTAACCCAGAGGATTGGAAATATCAGATGAA	6014	CC	a spot in the array. The method is useful for measuring the expression of		
Qy	976	TTCAATCTCAGCCGGTGCAGCTGGCTCTACCATGGTGGACCCGCTGGAGGCC	1035	CC	genes in a first Bacillus cell relative to expression of the same genes		
Db	6015	ATGGCTGCAGTTCTGTGAATGCTCTGAATCTAACTCCCTAAACATCATG	6074	CC	in one or more second Bacillus cells. The method is useful for monitoring		
Qy	1036	TCGGCCCTGGCAGITCACACACCGGGATATCAGCACTGGCCACGACTCGATG	1095	CC	global expression of several genes from a Bacillus cell, discovering new		
Db	6075	AAGGAGCTGGAGATAGATGCCGATTATCGTTAGTGGCAGCACGATG	6134	CC	genes, identifying possible functions of unknown open reading frames and		
Qy	1096	ATGCCCTCAGGTCACGGAGCTGGGAATGGTTCTCAAGCCGTTGGCCAGGCAC	1155	CC	monitoring gene copy number variation and stability. Monitoring changes		
Db	6135	TTAGGCGAAATAACAGAAATAGGAATGGTGTGGCTTAATGGTATAAGGCCAC	6194	CC	in expression of genes may be used to provide a representation of the way		
Qy	1156	GTTCCCAGAAGAATGGACCGATTTCGATGACCTTCGAAAGGAACCTGAGGT	1205	CC	in which Bacillus cells adapt to changes in culture conditions,		
Db	6195	TCACCAAAAGAGTGGGGAAATAAGGAAATGGTGTGGCTTAATGGTATAAGGT	6244	CC	environmental stress or other physiological provocation. Extensive follow		
Qy	1156	RESULT 14		CC	-up characterisation is unnecessary, when one spot on an array equals one		
Db	6195	ABK73017		CC	gene or one open reading frame, since sequence information is available.		
Qy	1156	ID ABK73017 standard; DNA; 1218 BP.		CC	This sequence represents a genomic sequence tag (GST) used in the method		
Db	6195	XX		CC	of the invention. Note: The sequence data for this patent did not form		
Qy	1156	RESULT 14		CC	part of the printed specification, but was obtained in electronic format		
Db	6195	ABK73017		CC	directly from WIPO at ftp://ftplib.wipo.int/pub/published_pct_sequences		
Qy	1156	RESULT 14		CC	XX		
Db	6195	ABK73017		CC	Sequence 1218 BP; 305 A; 268 C; 370 G; 275 T; 0 U; 0 Other;		
Qy	1156	RESULT 14		CC	Query Match 12.1%; Score 150.2%; DB 6; Length 1218;		
Db	6195	ABK73017		CC	Best Local Similarity 47.3%; Pred. No. 3.2e-26;		
Qy	1156	RESULT 14		CC	Mismatches 543; Indels 6; Gaps 1;		
Db	6195	ABK73017		CC	Matches 492; Conservative 0;		
Qy	1156	RESULT 14		CC	QY 180 CGGAAACATCATGGGCCGACGTGATGGCAAGGCACTGATGGCTGGCTGGCG 239		
Db	6195	ABK73017		CC	DB 167 CGGAAATCTGTTGGAGAATAGAAGGTACAGAGAACCGTACGGGATCTTGACCGG 226		
Qy	1156	RESULT 14		CC	QY 240 TTCAACACTCGATTCTGTCGGCTGGCACTGCGGACTGTCGGCTGGTGTG 299		
Db	6195	ABK73017		CC	DB 227 TTACACATTGATAACGGTCATCAACGGGAAATTGACGGAGGTACGGGATTAGGC 286		
Qy	1156	RESULT 14		CC	QY 300 CGCCCTTGAGGTGCCCCGGTGTGGAAACGGCTACGTGAAATGGCATCCATTG 359		
Db	6195	ABK73017		CC	DB 287 GAGCCTTTGrrGCAAAACACCTCTGGCACGGCAACGTACGGCAAGGGCGA 346		

Qy	360	GTTCATCGGATCGGGAGGAACGGGGCCGTTCAAGCAGTGGCATGTTGGGGCG 419	XX	Bacillus subtilis.
Db	347	AGTCGTGCTCTGTGGAGGAAGAGGGAAAGCGGTTCCCTTGACTTTGGGATCGGG 406	OS	
Qy	420	GGCCATGGGGTGGTCGCGAACAGGGAAACTGGACTCTTGGTGTATGGAGATGGGT 479	XX	Location/Qualifiers
Db	407	GAATTAAACGGGCTTGTGTAAGACTGTCAAAAGACGCCGGTGTGACGATCGGGACGGGT 466	FH	112. .135_0
			FT	
			FT	
			PN	W09934003-A2.
Qy	480	GTCGGTAGGGCAGGGCTACTGCCCTTGAGGGAAACTGGCAGGGTGCAGGC 539	XX	
Db	467	ATCGATTGAGCAGGGATGAAAGATTGGGTT-----CGGAAAACGGGGTACAGGGC 520	PD	08-JUL-1999.
Qy	540	CCGCTCGGGGGGACCTGCGCTTATCGAACACTACACATTGAAACAAGGACCGATCCT 599	XX	17-DEC-1998;
Db	521	CCCGTTAGGGTGTAGTGAAAGATTGGCTTATCGAGCTTCATATTGAAACAGGGGAGCATTCT 580	PF	98WO-US027020.
Qy	600	CGAGCAGGAGCAAATAAGAGATGGAGTGTAACTCCATCGTGGGTTCGGCATTGCG 659	XX	
Db	581	TGAGGAGGGCGTCAAATGGCATTGTCAGGGACATCGTGGACAAAGACGGTTAAC 640	XX	30-DEC-1997;
Qy	660	GGTTGCCTAAAGGGAGAACGCCACGGGAAACCCCCATGCACCTGGCCAGGA 719	XX	97GB-00027464.
Db	641	GGTTATGGTCAAGGGAGAACCATGGGACTGGGAGGTCAACCGAGATCGC 779	XX	(GEMV) GENENCOR INT INC.
Qy	720	TGGCTGGTACCCGGCTCTCATGGTGAGGGAGGTCAACCGGGTTGTCAACCGAGATCGC 760	XX	
Db	701	TGGCTGGCCGCTCTCGGCAATTGCGATTGTCATGGAAACAAAGCTAAAGCTGGCTA 760	XX	
Qy	780	CGATGGCACAGTGGCTACAGTGGGCCACCTCACAGTGGCCCTCCAGGAAACCAGGT 839	XX	This DNA sequence codes for a novel metalloprotease (MP), designated YurH (see AAY06396), of <i>Bacillus subtilis</i> . It was identified via a BLAST search of <i>B. subtilis</i> genomic DNA, and shows homology to the MP succinyl-diaminopimelate desuccinylase of <i>Escherichia coli</i> . An expression vector including the present sequence and a host cell comprising the vector are claimed. Also claimed are a cleaning composition, an animal feed and a composition for the treatment of a textile, all comprising YurH. Gram positive microorganisms having a mutation or deletion of all or part of YurH DNA are used as host cells for expression of a homologous or heterologous protein, such as a hormone, growth factor, cytokine or enzyme, especially a protease, carbohydrolase, lipase, isomerase, oxidase, reductase, transferase, kinase or phosphatase (all claimed). Also claimed is a method for detecting a Gram positive microorganism MP using a probe comprising all or part of the YurH DNA
Db	761	TCCGCTTTAACAGCAACGGTGGGAGGATTGAAGCAAAGCCGAATGTGCCGAATGTCAT 820	XX	
Qy	840	CCGGGGAGGTGACTTCACACTGGACTCTGGCTTCCAGTGGCCATGAGGAGTGGCTCCGGT 899	XX	
Db	821	ATCAGGAGAACGGTGTCTCTTGATCTCGGCCATCATGATGGGGTGTGGACCG 880	XX	
Qy	900	GCTGATGACCGCATCTCGGTATGGTGGGAGGTGGCTCCAGGGGGTGTGGCTGC 959	XX	
Db	881	GTATTGGGACGATATTGGTTATTCACAGAGCTGGGGGAATGGAACTGGTCAGAT 940	XX	
Qy	960	CGATGGGATGAATTTCAACTCAGCCGGTGTGACGTGGCTCTACCATGGTGGACGC 1019	XX	
Db	941	CACAGTGTAAACAGACGACGGATGCGGAAGGGTGTGACGGATGGACGGGATTGATTCAACT 1000	XX	
Qy	1020	CGTTCGGAAGGGCCTCGGCTTCAAGCAGCTGGGATATCAGGAGTGGGCGGG 1079	XX	Sequence 1450 BP; 397 A; 310 C; 415 G; 328 T; 0 U; 0 Other;
Db	1001	CTCAAAACAGTCGGCCGAGGATCTTGACGTATCCTATCAGGAGATGGTCAGGGGGGG 1060	XX	Query Match 11.7%; Score 145.2%; DB 2; Length 1450;
Qy	1080	CCACGACTCGATGTTCATCGGCCAGGTACGGCACGTGGGAATGGTTTCTGTTCCAAGCG 1139	XX	Best Local Similarity 46.4%; Pred. No. 5.2e-25;
Db	1061	GCACGATGCCAGGTTTCGGCCGATTCGGCGACGTCACTGCTGTTCCGAGCG 1120	XX	Mismatches 588; Indels 6; Gaps 1;
Qy	1140	TGCTGGGGAGCCACGTTCCGAAGAAATGGACCTTTCGATGACCTTCCGAAGAAC 1199	XX	Matches 514; Conservative 0;
Db	1121	GAAGGGCATCAGCATTGGGGTGTGAGCTTGCAGCGGGGT 1180	XX	
Qy	1200	TGAGGTGTCCCTCGGGTAAT 1220	XX	
Db	1181	TGAGCTCTGACTCATGTT 1201	XX	
RESULT 15			Db	
ID	AAX59333	standard; DNA; 1450 BP.	Db	
XX			Db	
AC	AAX59333;		Db	
XX			Db	
DT	20-SEP-1999	(First entry)	Db	
XX			Db	
DE	Bacillus subtilis metalloprotease YurH DNA.		Db	
XX			Db	
RW	Metalloprotease; protease; YurH; determinant; surfactant; cleaning; textile; feedstuff; animal feed; host cell; ds.		Db	
KW	425 TTGCAAGGGTTGGTGGCGACAGGGAACTGGACTCTTGGTGTAGGAGTGGTGTCCCG 484		Qy	

Db	542	TGACGGGTRTTTCAAGAACAAAGATGCCAAAGGCCAAGAGCAGTCGGAGTCCTCCC	601
Qy	485	TTAGGCAGGGCGCTACTGCCCTCGGTTGAAGCCGGGAACCTGCAGGCTGGCTCC	544
Db	602	TCCAAACAGCGATGCATGAGAGGGTTTGCAAAAGGTGTTCAGTCAGCA-----T	655
Qy	545	CGCGGGGACCTGCGTGCCTTATCGAACTACACATTGAAACAAGGCCATCCCTCGAGC	604
Db	656	ATAGGACAGATACTAGGGGTTGTTGAGCTGCATATTGAAACGGGAAGACGTTGAAA	715
Qy	605	AGGCAAATAGAATCGGAAGTTGTAACCTCCATGTTGGCGATTGGCGATTTGGGGTTG	664
Db	716	TGTCAGGGGGATCTCGGCACTCGTGAAGTATTGGGGGAGAGACGATATCTCGTCA	775
Qy	665	CGTCAAAAGCAGAACACCGCCACAACCCCCATGCAACCTGGCAGGATGGCG	724
Db	776	CGCTCGAAGGAAATGCAATCACGAGGAACCACCTCCATGAAATGGCGAAGGCCG	835
Qy	725	TGGTACCCGGCTCATGGTGAGGGAGGTCAACCGGTTGCAACGAGATGCCGATG	784
Db	836	TGGAGCCAGCAGGGTCAATTCATGAGCTGGCAATGGCCAATGTCATGAGCTGGG	895
Qy	785	GCACAGTGGTACCGTGGCACCTCACACTGGACCTGCTGGCCCGGTGAGGCAAC	844
Db	896	AGCTCCGGTGTGACATGGGAAATAACGGCAAGGCCAATGTAGCCAATACCGG	955
Qy	845	GGAGGGTGGACTTCAACTGGACCTGCTGGTTCTCCGCATGAGGAGTCGGCTG	904
Db	956	GCGCGTCCAGTTCAATCGATATTGCCATCAGCATCGCATGCTGGTGGAAACAGT	1015
Qy	905	TGACCGGCACTCGGTCACTGGTGCGGAGGTGCGCTCCAGGCCGGTGTGCGATG	964
Db	1016	ATCAAGACATGGTGTGCTTGTCAACGGCATTTGCCTACAAAAGGAATTGGCTGTGA	1075
Qy	965	TGGATGAATTTCATCTAGCCGGTCAACACACCGGGATATCGGCAGTCGGCTC	1024
Db	1076	TTGATGAATATATGGGATAGAGCCTGTCGCGATGGACGAAAGGCTGAAGGCT	1135
Qy	1025	GCGAAGGGCCTGGCTTGCAGTTCAACACACCGGGATATCGGCAGTCGGCTC	1084
Db	1136	TTGAAACAGCATTAGAAAAGGCTTCAGCTGTGCGATGGACGAAATGGTGAAGGG	1195
Qy	1085	ACTCGATGTTCATGCCCAAGTCAGGTGGACGTCGGAAATGGTTTCGTTCCAAG	1144
Db	1196	ACGGGAAATGATGGAAAGGGCTATCTGCTGTTGATGCTGTTGAGCCGAGGCG	1255
Qy	1145	GCGGAGGCCACGTTCGGAAAGAATGGAC	1172
Db	1256	GGTCAAGCCACTCACCGAAGGAATAAC	1283

Search completed: May 4, 2004, 13:34:47
 Job time : 624 secs

This Page Blank (uspto)

Total number of hits satisfying chosen parameters: 1365418

minimum DB seq length: 0
maximum DB seq length: 2000000000

post-processing: Minimum Match 0%

Listing first 45 summaries

```

issued_fatents_na : *
acdbase :
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfile1.seq: *

```

No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	123.9	100.0	1239	4	US-09-285-055-1
2	214.8	17.3	1488	4	US-09-252-991A-6822
C	214.8	17.3	1698	4	US-09-252-991A-6923
3	210.8	17.0	1380	4	US-09-489-039A-2442
4	145.2	11.7	1450	4	US-09-554-998A-1
5	122.6	9.9	1278	4	US-09-134-000C-2571
6	121.2	9.8	1830121	4	US-09-557-884-1
7	121.2	9.8	1830121	4	US-09-643-990A-1
8	104.8	8.5	924	4	US-09-489-039A-4854
9	104.8	8.5	924	4	US-09-543-681A-1026
10	97	7.8	1290	4	Sequence 1026, AP
C	93.4	7.5	480	4	Sequence 2592, AP
C	93.2	7.5	525	4	Sequence 6967, AP
C	84	6.8	564	4	Sequence 6968, AP
C	76.8	6.2	339	4	Sequence 2583, AP
C	56.6	4.6	405	4	Sequence 4973, AP
C	52.8	4.3	300	4	Sequence 2448, AP
C	48.8	3.9	381	4	Sequence 15810, A
C	48.8	3.9	537	4	Sequence 15911, A
C	48.8	3.9	870	4	Sequence 15882, A
C	48.8	3.9	1317	4	Sequence 15836, A
C	48.8	3.9	2073	4	Sequence 15779, A
C	48.6	3.9	1104	4	Sequence 13171, A
C	48.6	3.9	1725	4	Sequence 12780, A
C	48.2	3.9	1425	4	Sequence 217, App
C	48.2	3.9	3114	4	Sequence 225, App
C	48.2	3.9	47981	4	Sequence 1, Appli

65418

No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description	
Sequence 1, Appli	
Sequence 6822, Ap	
Sequence 6923, Ap	
Sequence 2442, Ap	
Sequence 1, Appli	
Sequence 2571, Ap	
Sequence 1, Appli	
Sequence 4854, Ap	
Sequence 1026, Ap	
Sequence 2592, Ap	
Sequence 6967, Ap	
Sequence 6968, Ap	
Sequence 2583, Ap	
Sequence 4973, Ap	
Sequence 2448, Ap	
Sequence 15810, A	
Sequence 15911, A	
Sequence 15882, A	
Sequence 15836, A	
Sequence 15779, A	
Sequence 13171, A	
Sequence 12780, A	
Sequence 217, App	
Sequence 225, App	
Sequence 1, Appli	

ATTACHMENTS

; Sequence 1, Application US/09285055
; Patent No. 6352848
; GENERAL INFORMATION:
; APPLICANT: ALTENBUCHNER, JOSEF
; APPLICANT: MATTES, RALF
; APPLICANT: PIETZSCH, MARKUS
; APPLICANT: SYLDATK, CHRISTOPH
; APPLICANT: WIESE, ANJA
; APPLICANT: WILMS, BURKARD
; TITLE OF INVENTION: RECOMBINANT L-N-CARBAMOYLASE FROM ARTHROBACTER
; TITLE OF INVENTION: AURESCENS AND METHOD OF PRODUCING L-AMINO ACIDS

FILE REFERENCE: RECOMBINANT L-N-CARBAMOYLASE
CURRENT APPLICATION NUMBER: US/09/285,055
CURRENT FILING DATE: 1999-04-02

EARLIER FILING DATE : 11/29/04-02

Description	
Sequence 1, Appli	
Sequence 6822, Ap	
Sequence 6923, Ap	
Sequence 2442, Ap	
Sequence 1, Appli	
Sequence 2571, Ap	
Sequence 1, Appli	
Sequence 4854, Ap	
Sequence 1026, Ap	
Sequence 2592, Ap	
Sequence 6967, Ap	
Sequence 6968, Ap	
Sequence 2583, Ap	
Sequence 4973, Ap	
Sequence 2448, Ap	
Sequence 15810, A	
Sequence 15911, A	
Sequence 15882, A	
Sequence 15836, A	
Sequence 15779, A	
Sequence 13171, A	
Sequence 12780, A	
Sequence 217, App	
Sequence 225, App	
Sequence 1, Appli	

Qy	665 CGGTCAAAGGAGAACGGACACGCCAACCCATGCACCTGGCAGGAATGGCG 724	Qy	352 CCATTGAGTTCATCGGATCGTGGAGGAGGAAGGGGCCGTTTACGGAGTGGATGTTG 411
Db	776 CGCTCGAAGGAATGCAATCAGCAGGAACCCTCCATGAAATGGCGAAGGCCGC 835	Db	397 TCTTTAGAAGTGCTTCCATGGGTGAAGAAGGGATCGTGTTC 456
Qy	725 TGGTACCCGGCTCTCATGGTGGAGGTCAACGGAGATCGCCGATG 784	Qy	412 GGCGCCGGCCATTGCAGGGTGGTGGCCGACAGGGAACTGGACTCTTTGGTTGATGAG 471
Db	836 TCGCAGCCAGCAGGGTATCATTCAATGGCTGGATGAGCTGGGATG 895	Db	457 GGAAGCAAAAATGTCGTTGGTAAGAAGTAGTGTGATACTCTGACAAAC 516
Qy	785 GCACAGTGGCTACGTTGGCACCTCACAGTGGCCCCGGTGGAGGAACACGGTCCGG 844	Qy	472 GATGGAGTGTCCGTTAGGCAGGGGGCTACTGCCTTGGCTGAAGCGGGAACTGCAAG 531
Db	896 AGCTCCGCTGACATGGGAAAATAACGGCAGGCCAATGTAGCCAAATGCTACACGG 955	Db	517 GAGGTAAGAAAGTTGTTGATGAGATGCATGCCAAAGGCTTGTGAAAGAGGCA 576
Qy	845 GGGAGGTGGACTTACACTGGACCTGCTGCGTTCCTACAGTGGCTGGCTGATG 904	Qy	532 GCTGAGGCCCGCTCGGGGAACTCTGGTGGCTGCTGGTCAAGTGGCAACAGGA 591
Db	956 GCGCGTCCAGTTCAATGATATTGCCATCAGCATCAGCATGGCAACAGTTTC 1015	Db	577 ACTAGAC-----GTGAGGACATTGAAGGATTGGCAATTGNAACAAGGC 627
Qy	905 TCGACCGCATCTGGTCATGGTGGGAGGTGGCTGGCTGGCGATG 964	Qy	592 CCGATCCTCGAGCAGAGCCAATAGAGATGGAGTTGTAACCTCCATCGTTGGCGTTCGC 651
Db	1016 ATCAAGACATGGTGTCTTGAATCAACGGCATTTGCTACAAAAGGAATTGGCTGTGA 1075	Db	628 AATATTAGAGAACGAAAGCTCCAAGTGGTGTGTTCATAGTATTGTAAGGCAACGT 687
Qy	965 TGGATGAAATTTCAAATCTCAGCCGGTGCAGCTGGCTCCTACCATGGTGGACGCCGTT 1024	Qy	652 GCATTGGGGTTGCCGTCAAGGCAGAACCGGACCCACGGCACAACCCCATTGCACCTG 711
Db	1076 TTGATGAAATATGGGATAGAGCCTGCGCATGGACGAAGGCTGAAGGGCTT 1135	Db	688 CGCTACACGGTTAACCTAAAGGGCAAGCCAAATCATGAGGTTACACTGAGGTTAC 747
Qy	1025 GCGAAGGGCCCTGGCTGGAGTTCACACACGGGATATCACAGCAGTGGGGGGCACG 1084	Qy	712 CGCCAGGATGGCGTGTACCCGGCTCTCATGGTGGAGGTCAACCGGTTCGTCAAC 771
Db	1136 TTGAAACAGCATTAGAAAAGGGTTCAAGCTGTGAGGAATGGTGAAGGGCATG 1195	Db	748 CGTCATGATGGGTACGGTTGGCTAAATTGGTGCAGAAGGCCATTGATCGTGGAAAC 807
Qy	1085 ACTCGATGTTCATGCCCGTGGAGTCACGGACGTCGGATGGTTTCAGCCGTGCTG 1144	Qy	772 GAGATGCCCGATGGCACAGTGGTACCGTTGGCCACCTCACAGTGGCCCCGGTGGAGGC 831
Db	1196 ACGGCAAATGATGGAAAGGGCTATCCTGCTRTGATGCTGTTGCGAGGCCGGCG 1255	Db	808 GAAATTGGGGATTCGGTTAGTTTAACATTGGCAAAGTCATTCCGAAACCCATAACAGTC 867
Qy	1145 GCGGGAGCCACGTTCGGAAAGGAATGGAC 1172	Qy	832 AACCAAGGTCCCGGGAGGTGGACTTCACACTGGACCTGCGTTCCGCATGAGGAGTGC 891
Db	1256 GCGTCAGCCACTCACCGAAGGAATATAAC 1283	Db	868 AACGTGGTGTGCCCTGGAAAGTGGATTACATTGACTGCGGTACATACCGACGCTGGTTT 927
Qy	RESULT 6		
US-09-134-000C-2571	Query March 9.9%; Score 122.6; DB 4; Length 1278;		
; Sequence 2571, Application US/09134000C	Best Local Similarity 45.7%; Pred. No. 3.3e-22;		
; Patent No. 6617156	Matches 474; Conservative 0; Mismatches 554; Indels 9; Gaps 1;		
; GENERAL INFORMATION:	Db 1228 GAAGGTTCAAGCTTT 1244		
; APPLICANT: Lynn Doucette-Stamm et al.	RESULT 7		
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO	US-09-557-894-1		
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS	; Sequence 1, Application US/09557884		
; FILE REFERENCE: 032796-032	; Patent No. 6506581		
; CURRENT APPLICATION NUMBER: US/09/134, 000C	; GENERAL INFORMATION:		
; CURRENT FILING DATE: 1998-08-13	; APPLICANT: Fleischmann et al.		
; PRIOR APPLICATION NUMBER: US 60/055, 778	; TITLE OF INVENTION: The Nucleotide sequence of		
; PRIOR FILING DATE: 1997-08-15	the Haemophilus influenzae Rd Genome, Fragments		
; NUMBER OF SEQ ID NOS: 6812	Thereof, and Uses Thereof		
; SOFTWARE: Patentin version 3.1	; NUMBER OF SEQUENCES: 1		
; SEQ ID NO: 2571	; CORRESPONDENCE ADDRESS:		
; LENGTH: 1278			
; TYPE: DNA			
; ORGANISM: Enterococcus faecalis			
US-09-134-000C-2571	Qy 1192 AAAGGAACCTGAGGTTGT 1208		
;	Qy 172 GACGCCACTCGGAACATCATGCCGAGCTGAGGACTGTCCGGAGGTCTCCTGCGATC 231		
;	Db 217 GACGAAGTGGGAAATCTGTTGAGTGGGACCTTAGTTAGATGGTCATAATGGT 276		
;	Qy 232 GCGGTGGTTCACACTTCGATTCGTTCTGTCGAAACGGGGCTACGTGAATCGGCAT 291		
;	Db 277 TAACTGGTCACATATGCCACAGTAGTTAATGGTGGGACCTTAGATGGTCATAATGGT 336		
;	Qy 292 GTGGTGTGGCCCTTGAGGCTGCCGGTGTGGAGAACGGCTACGTGAATCGGCAT 351		
;	Db 337 GTCATTCGCTCCATGACAGCTGTTCAATATTAGACAATATGGGAAACACGGCGT 396		

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: MD

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: 3 1/2 inch diskette

COMPUTER: Dell 1 Pentium

OPERATING SYSTEM: MS DOS v6.22

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/557,884

FILING DATE: 25-Apr-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

NAME: Michelle S. Marks

REGISTRATION NUMBER: 41,971

REFERENCE/DOCKET NUMBER: PB186P3

TELEPHONE: 301-309-8504

TELEFAX: 301-309-8439

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1830121 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-557-884-1

Query Match

Best Local Similarity

Matches

Conservative

0; Mismatches

673; Indels

0; Gaps

0;

Qy 568 ATCGAACTACACATTGAAACAAGGACCGATCTCGAGGAGCAAATAGAGATCGGAGTT 627
 Db 609200 TTGAACTCATATAAGAACGGCCCTCGTTAGAGAATGAAGAAAACAATAGGTGTT 609259
 Qy 628 GTAACCTCATCGTGGCGCAATTGGGGTTGCCGCTCATCGTGAATTAAAGGACAAGGGATCAT 687
 Db 609260 GTGACAGGTATTGCTGCCGATTCGTGCAATTGTGTTAAAGGACAAGGGATCAT 609319
 Qy 688 GCGGCCAACCCCCATGCACCTGGCCAGGATGGCTGGCTCATCGTGGCTACGGTGGCCAC 807
 Db 609320 TCAGGAGAACAGCAATGCATTATCGTCATGATGCATTGTAGGGATCTGAGTTATCA 609379
 Qy 748 ACGGAGGTCAACGGGTTCGTCAACGAGATGCCGATGCCACAGTGCTGGACTCAGCTGGAC 867
 Db 609380 CTTGCTATTGAGCGAGCTGCTATCAAGCTGGACATTCTACAGTAGCTACAGTAGTTAAT 609439
 Qy 808 CTACAGTGGCCCCGGTGGAGGCNAACCAGGTCACCGGTTCCGGATGCCGATCTCGGTGTCATGGTC 927
 Db 609440 ATTACAGCTAACAGGAGTAATGAATGTTGCTGCCAGGATATTGGAATTATTAGTAGAT 609499
 Qy 868 CTGCGTTCTCCGCATGGAGGTGGCTGGCTGGCTGATGCCGACCGCATCTCGTCAGTC 928
 Db 609500 ATTGAGGTACATGTACAAGCTAGAGATCTGTATTGAATTATTACAAGGAGAAATT 609559
 Qy 928 GCGGAGGTGCCTCCAGGCCATGCCGATGGATGAAATTTCATCTCAGTC 987
 Db 609560 AGTAAAGTTTCAGAAAAAGGGATTATTAAATTGAGATTACAACCTTCTGTTGAATT 609619
 Qy 988 CGGGTGCAGCTGGCTCCACCATGGGGACGCCGTTGCCGAAGGGCCTGGCTGCAG 1047
 Db 609620 CCAATAATTATTACCTGAAATAATGGTAACACAAATAGCTGAAACATTCCCTTGGT 609679
 Qy 1048 TTCACACACCGGGATATCAGGAGTGGGGGGGCCACGACTCGATGTTCATCGGCCAGGGC 1107
 Db 609680 TATTCTTACGAAATTATGCCAAGTGGTGCAGCATATGGCAATATGCAACTT 609739
 Qy 1108 ACGGAGCTGGGAATGGTTTCGTTCCAAGGCCGGTGGCTGCCAGGTCACGTTCCCGAAAGAA 1167
 Db 609740 TGCCCCAACAGGTATGTTTATCCATCTCATTGGAAATTAGTCATAACCCCTCTGAA 609799
 Qy 1168 TGGACCGATTTCGATGACCTTCGCAAAAGGAACACTGAGGTGTCCTCAGGGTAATGAAAGG 1225
 Db 609800 TTAACTGTGAAAGATATAGGGCAGGAAATTACAAAAGTTATATTGAA 609857

RESULT 8

US-09-643-990A-1

; Sequence 1, Application US/09643990A

; Patent No. 6528289

; GENERAL INFORMATION:

; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:
 MEDIUM TYPE: 3 1/2 inch diskette
 COMPUTER: Dell Pentium
 OPERATING SYSTEM: MS DOS v6.22
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:

APPLICATION NUMBER:	US/09/643,900
FILING DATE:	23 -Aug-2000
CLASSIFICATION:	<Unknown>
PRIOR APPLICATION DATA:	
APPLICATION NUMBER:	08/487,429
FILING DATE:	1995-06-07
APPLICATION NUMBER:	08/426,787
FILING DATE:	1995-04-21
ATTORNEY/AGENT INFORMATION:	
NAME:	Kenley K. Hoover
REGISTRATION NUMBER:	40,302
REFERENCE/DOCKET NUMBER:	PB186P1C1
TELECOMMUNICATION INFORMATION:	
TELEPHONE:	301-610-5790
TELEFAX:	310-309-8439
INFORMATION FOR SEQ ID NO: 1:	
SEQUENCE CHARACTERISTICS:	
LENGTH:	1830121 base pairs
TYPE:	nucleic acid
STRANDEDNESS:	double
TOPOLOGY:	linear
SEQUENCE DESCRIPTION:	SEQ ID NO: 1:
S-09-643-990A-1	
Query Match	9.88 ; Score 121.2 ; DB 4 ; Length 1830121 ;
Best Local Similarity	43.88 ; Pred. No. 7.1e-21 ;
Matches	525 ; Conservative 0 ; Mismatches 673 ; Indels 0 ; Gaps 0 ;
b	28 CGCATGAGAAAGAGATCCGGGAGCTCTCCGGTCTCGGCAGAAGGGCCCGGTAC 87 608660 CGTGTCAAAATTAAATTGAAAACACTGGCTTTATTTCATCAGTACCAAACGAGCTGACT 608719
y	88 CGGCTGACCTACACTCCAGAGCATGCCGCCGGGGAAACGCTCATGGCTATGAAA 147 608720 CGCTTAGCTTTACAGAGGAGATGAAAAGGCCATATAATGATTATGAAATTGTAAA 608779
b	148 GGGCCGCCTGAGCGCTTCTGGATGGCAACTCGGAACATCATCGGCCGACGTGAAGGGC 207 608780 GAATATGATTGTCTATTCTGTCAGATTCAATTGGAAATCTTTTATTCTGTAAGGCAGGT 608839
b	208 ACTGATCCGGAGCTTCTGGATGGCTCAGACTTCGATTCTGTCGGAAACGGC 267 608840 AAAGAAAGATTTTACCTGCGATTGCAATTGGAAATCTGGATCACATTGATACTGTTGTGAATGCT 608899
y	268 GGGATGTTGATGGCACTGCAGGGCTTGGTGTGCCCTTGAGGCTGCCCCGGGTGATGCTG 327 608900 GGTAATTGATGGTCCTTGGATCTGTTGCTGGTTGAAATTCTTGTCAATTGTTG 608959
b	328 GAGAACGGCTACGTGAATGGCATTCGATTGAGTTCATCGGATCGTGGAGGAAGGG 387 608960 AACAGAAATTATCAGACTCGTTAGAAATTGATAATTGTGAAATTGTA 609019
y	388 GCCCGCTCAGCAGTGGCATGTTGGCCATTGCAAGGGCCATTGCAAGGGACAGG 447 609020 AGTAGATTAAATTGCTACATTGGTAGTAAGTTATGTGTGGCATAGTAATCAAGAA 609079
b	448 GAACTGGACTCTTGGTTGATGGAGATGGGTCTGGCAGGGCTACTGCCTTC 507 609080 AAAATAAGTTCAATTACGTGATAAACAGGAAAGGTTTATCAGAAAGCTATGGCTGAAGTA 609139
b	508 GGCTTGAAGCCGGCGAACACTGCAGGGCTGCCGCTCGGGGACCTCGGTGTTT 567 609140 GGAATGAAATTTAATTGGTTAATCAAGCAAAGGCCCTCGTTAGAGATGAAGGAATTAAATGTT 609199
b	568 ATCGAACTACACATTGAACAAGGACCGATCCCTCGAGGAGGAAATAGAGATCGGAGTT 627 609200 TTGTAACCTTCATAGAAAGGCCCTCGTTAGAGATGAAGGAATAACAGGGATCAT 609259
y	448 GAACTGGACTCTTGGTTGATGGAGATGGGTCTGGCAGGGCTACTGCCTTC 507 609080 AAAATAAGTTCAATTACGTGATAAACAGGAAAGGTTTATCAGAAAGCTATGGCTGAAGTA 609139
b	508 GGCTTGAAGCCGGCGAACACTGCAGGGCTGCCGCTCGGGGACCTCGGTGTTT 567 609140 GGAATGAAATTTAATTGGTTAATCAAGCAAAGGCCCTCGTTAGAGATGAAGGAATTAAATGTT 609199
b	568 ATCGAACTACACATTGAACAAGGACCGATCCCTCGAGGAGGAAATAGAGATCGGAGTT 627 609200 TTGTAACCTTCATAGAAAGGCCCTCGTTAGAGATGAAGGAATAACAGGGATCAT 609259
Query Match	8.58 ; Score 104.8 ; DB 4 ; Length 924 ;
Best Local Similarity	47.58 ; Pred. No. 1.1e-17 ;
Matches	310 ; Conservative 0 ; Mismatches 342 ; Indels 0 ; Gaps 0 ;
Qy	569 TCGAACTACACATTGAACAAGGACCGATCCCTCGAGGAGGAAATAGAGATCGGAGTTG 628 Db 245 TTGAGCTGCATATCGAACAGGGACATTCTGTGATAAGAGCAGATTGATTTGGCTGG 304
Qy	629 TAACCTCCATTGTTGGCTGGCATTGGGGATGGCTGCGCTCAGGGTCAAGGCAAGCAG 688 Db 305 TTACGGGGTACAGGGTATTCTGTGGCAGGAATTACCCCTGAGAGCTGATGGCTGG 364
Qy	689 CCGGCACAAACCCCATGACCTGGCTGGCAGGATGCCGCTGGTACCCGGCTCTCATGGTGA 748 Db 365 CGGGTACTACCCCAATTGTCATAGAGACGTGATGGCTGGCTGGAAATTCGCTG 424
Qy	749 GGGAGGTCAACCGGTTCTGTCACCGAGATCGGCCAGGGCTACCGGCCACC 808

; TYPE: DNA
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-4973

Qy 506 TCGG 509 **Db** 64 AGGAGACCCCTGCCAAGGCCGATGCCAGGGTAGGTAGGGCTGGACGCCA 5
Qy 4 TCGG 1 **Db** Sequence 2583, Application US/09489039A
Qy Patent NO. 6610836
Qy ; GENERAL INFORMATION:
Qy ; APPLICANT: Gary Breton et. al
Qy ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
Qy ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
Qy ; FILE REFERENCE: 2709.2004001
Qy ; CURRENT APPLICATION NUMBER: US/09/489,039A
Qy ; CURRENT FILING DATE: 2000-01-27
Qy ; PRIOR APPLICATION NUMBER: US 60/117,747
Qy ; PRIOR FILING DATE: 1999-01-29
Qy ; NUMBER OF SEQ ID NOS: 14342
Qy ; SEQ ID NO 2583
Qy ; LENGTH: 339
Qy ; TYPE: DNA
Qy ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-2583

Qy ; Query Match 4.6%; Score 56.6; DB 4; Length 405;
Qy ; Best Local Similarity 55.2%; Pred. No. 2.1e-05;
Qy ; Matches 132; Conservative 0; Mismatches 104; Indels 3; Gaps 1;

Qy 72 AGGCCCGGTACCCGGCTGACCTACACTCCAGAGCATGCCGCCGGAAACGGCT 131
Db 102 AGGGGGGGTATGCCCTGGCAGCGACTGCCGAAGATAAGGGACTTTGT 161

Qy 132 CATGCCGCTATGAAAGGGCCCTTGAGCGTTCTGAGCTCGTGAAGAACATCAT 191
Db 162 TGTAGCCCCATGAAGGGCGCTGGGAAATGTCAC 221

Qy 192 CGGGGACGTGAAGGCCACTGATCCGGAGCTTCTGGATCAGCTTCGA 251
Db 222 CGCGTTATCACGGCGAGGAGACGTTGCCGATGGTGTATGGCAATACGGCAAAATACGGCAAT 278

Search completed: May 4, 2004, 16:04:11
 Job time : 141 secs

Qy ; Query Match 6.2%; Score 76.8; DB 4; Length 339;
Qy ; Best Local Similarity 52.1%; Pred. No. 1.3e-10;
Qy ; Matches 171; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

Qy 73 GGCCCGGTACCCGGCTGACCTACACTCCAGAGCATGCCGCCGGAAACGGCTC 132
Db 328 GCCGACCCCTGAACCCGGCTATCTGTGCCCGAGCATCTGCAGGCCAACAGCTGGTG 269

Qy 133 ATTGGGGCTATGAAAGGGCACTGATCCGGAGCTTCTGGCTGAAGAACATCATC 192
Db 268 GGGCAGTGGATGCCGGCATGATGGTGGCTGGAGACAGCGTGGCAACATCTGC 209

Qy 193 GGCCGACGTGAAGGGCACTGATCCGGAGCTTCTGGCTGAAGAACATCATC 192
Db 208 GAACGTTATGAAGGACAACAGGAGGAAACAGGAGCAACAGGAGCAACATCTGC 149

Qy 253 TCTGTCGGAAACGGGGATGTTGATGGCACTGCAAGGGACTCGCCCTTGAGGCT 312
Db 148 ACCGTTGGCAACGGGGACGGCTACGGCATGCTGGGTGCTGGGGATGGGTG 89

Qy 313 GGGGGGTGATGCTGGAGAACGGCTACGTGAATCGGCATCCATTGAGTTCATCGCGATC 372
Db 88 GTCCAGGGTCTGCACCAAGCATGGGCTGGCGAACGGCAATCGAGATCGGGCTTT 29

Qy 373 GTGGAGGAGGAAGGGGGCGCTTCAGCA 400

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using BW model

Run on: May 4, 2004, 15:00:24 ; Search time 618 Seconds
 (without alignments)
 9068.887 Million cell updates/sec

Title: US-10-045-063-1

Perfect score: 1239

Sequence: 1 gttgaccctgtcgaaaaagcgca.....tgaaggcacttgaccggtaa 1239

Scoring table: IDENTITY_NUC
 Gapop 10.0 , Gapext 1.0

Searched: 2936184 seq#, 2261732022 residues

Total number of hits satisfying chosen parameters: 58723368

Minimum DB Seq length: 0

Maximum DB Seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*

15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*

16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*

17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

ALIGNMENTS

RESULT 1
 US-09-916-501-3
 Sequence 3, Application US/09916501
 ; Patent No. US20020132848A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KRIMMER, Hans-Peter
 ; REICHERT, Dietmar
 ; APPLICANT: DRAUZ, Karlheinz
 ; APPLICANT: KLEMENT, Ingo
 ; APPLICANT: MAY, Oliver
 ; TITLE OF INVENTION: Process for the Preparation of Allysine Acetal
 ; FILE REFERENCE: 210740US-10757-9350-0-X
 ; CURRENT APPLICATION NUMBER: US/09/916,501
 ; CURRENT FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: Germany 100 37 115.9
 ; PRIOR FILING DATE: 2000-07-28
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 1239
 ; TYPE: DNA
 ; ORGANISM: Arthrobacter aurescens
 ; US-09-916-501-3

Pred. No. 18 is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1238	99.9	1239	9	US-09-916-501-3
2	1238	99.9	1239	15	US-10-156-761-3482
3	1192.4	96.2	1263	9	US-09-916-501-6
4	739.8	59.7	1239	9	US-09-950-772-5
5	739.8	59.7	1239	15	US-10-289-360-3
6	739.8	59.7	3343	9	US-09-950-772-7
7	739.8	59.7	3343	15	US-10-289-360-5
8	220.6	17.8	1163020	16	US-10-398-221-10
9	220.6	17.8	3011208	16	US-10-398-221-2058
c 10	158.6	12.8	3880	16	US-10-398-221-3817
11	150.2	12.1	1218	9	US-09-974-300-308
12	122.6	9.9	32768	9	US-09-070-927A-123
13	121.2	9.8	1830121	15	US-10-329-960-1
14	121.2	9.8	1830121	16	US-10-329-670-1

Db	1202	AGTTGTCCTCGGTAATGAAGGCACTTGACGGTAA 1239	
		RESULT 2	
	US-10-334-990-6	; Sequence 6, Application US/10334990	
		; Publication No. US20030175910A1	
		GENERAL INFORMATION:	
		APPLICANT: ALTENBUCHNER, JOSEF	
		APPLICANT: BOMMARIUS, ANDREAS	
		APPLICANT: MATTES, RALF	
		APPLICANT: SYLDATK, CHRISTOPH	
		APPLICANT: TISCHER, WILHELM	
		APPLICANT: WIESE, ANJA	
		APPLICANT: WILMS, BURKARD	
		TITLE OF INVENTION: WHOLE CELL CATALYST	
		FILE REFERENCE: 9350-0142-0	
		CURRENT APPLICATION NUMBER: US/10/334,990	
		CURRENT FILING DATE: 2003-01-02	
		PRIOR APPLICATION NUMBER: US/09/407,062	
		PRIOR FILING DATE: 1999-09-28	
		NUMBER OF SEQ ID NOS: 11	
		SOFTWARE: PatentIn version 3.1	
		SEQ ID NO 6	
		LENGTH: 1239	
		TYPE: DNA	
		ORGANISM: Arthrobacter aurescens	
		FEATURE:	
		NAME/KEY: CDS	
		LOCATION: (1) .. (1239)	
		OTHER INFORMATION:	
		US-10-334-990-6	
		Query Match 99.9%; Score 1238; DB 15; Length 1239;	
		Best Local Similarity 100.0%; Pred. No. 0;	
		Matches 1238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	2	TGACCCCTGCAGAAAAGGCCAACGGGGCATTCGAGAAAGAGATCCGGAGCTCTCCGGGT 61	
Db	2	TGACCCCTGCAGAAAAGGCCAACGGGGCATTCGAGAAAGAGATCCGGAGCTCTCCGGGT 61	
Qy	62	TCTCGGCAAGAGCCCCGGTTAACCGGGCTACACTCCAGAGCATGCCGCCGC 121	
Db	62	TCTCGGCAAGAGCCCCGGTTAACCGGGCTACACTCCAGAGCATGCCGCCGC 121	
Qy	122	GGAAAACGCTCATGGGCTATGAGCTATGAAAGGGCCCTTGAGCCTTGAAAGACGCACTG 181	
Db	122	GGAAAACGCTCATGGGCTATGAGCTATGAAAGGGCCCTTGAGCCTTGAAAGACGCACTG 181	
Qy	182	CAAACATCATGGCGGACGTGAAGGCACACTGATCGGGACTGTCCTGGATCGGGTGGCT 241	
Db	182	CAAACATCATGGCGGACGTGAAGGCACACTGATCGGGACTGTCCTGGATCGGGTGGCT 241	
Qy	242	CACACTTGATTCTGTGGCTCATGGGACTGTCCTGGATGGCACTGCAAGGGTGGCT 301	
Db	242	CACACTTGATTCTGTGGCTCATGGGACTGTCCTGGATGGCACTGCAAGGGTGGCT 301	
Qy	302	CCCTTGAGGCTGCCGGGTGATGCTGAATTCGCAATTGGACTGTCCTGGGGGGGG 361	
Db	302	CCCTTGAGGCTGCCGGGTGATGCTGAATTCGCAATTGGACTGTCCTGGGGGGGG 361	
Qy	362	TCATCGGATCGTGGCTCATGGGAGGAACGGCTACGTGAATTCGCAATTGGACTGTCCTGGGGGGGG 421	
Db	362	TCATCGGATCGTGGCTCATGGGAGGAACGGCTACGTGAATTCGCAATTGGACTGTCCTGGGGGGGG 421	
Qy	422	CCATTGCAAGGGCTACTGCCTCGGCTGAGCTGGCTGAGGATGGAGTGT 481	
Db	422	CCATTGCAAGGGCTACTGCCTCGGCTGAGCTGGCTGAGGATGGAGTGT 481	
Qy	482	CCGTTAGGCAGGGCTACTGCCTCGGCTGAGCTGGCTGAGGACTGAGGGCTGGAGCCC 541	
Db	482	CCGTTAGGCAGGGCTACTGCCTCGGCTGAGCTGGCTGAGGACTGAGGGCTGGAGCCC 541	

Qy	542	GCTCCGGGGGACCTGGAGTTGTAACCTCCATTATCGAAACTACATCGAACAGGACCGATCCTCG	601	; LOCATION: (25) . . (25)		
Db	542	GCTCCGGGGGACCTGGAGTTGTAACCTCCATTATCGAAACTACATCGAACAGGACCGATCCTCG	601	; OTHER INFORMATION: n=any nucleotide		
US-09-916-501-6				US-09-916-501-6		
Qy	602	AGCAGGAGAAATAAGAGATCGGAGTTGTAACCTCCATTATCGAAACTACATCGAACAGGACCGATCCTCG	661	Query Match 96.2%; Score 1192.4; DB 9; Length 1263;		
Db	602	AGCAGGAGAAATAAGAGATCGGAGTTGTAACCTCCATTATCGAAACTACATCGAACAGGACCGATCCTCG	661	Best Local Similarity 97.8%; Pred. No. 0;		
Qy	662	TTCGGTCAAGGCAGAACGGGACCAACCCCCCATGCACTTGCGGCCAGGATG	721	Matches 1208; Conservative 0; Mismatches 27; Indels 0; Gaps 0;		
Db	662	TTCGGTCAAGGCAGAACGGGACCAACCCCCCATGCACTTGCGGCCAGGATG	721	2 TGACCTGCGAAAGGCCAAGGGGCGCATTGAGAAAGAGATCCGGGAGCTCTCCGGGT	61	
Qy	722	CGCTGGTACCCGCCGCTCTCATGGTGAAGGAGATGCCG	781	2 TCTGGCAGAACGGCCGGCTTACCCGGCTGACACTCCAGAGCATGGCGGCC	121	
Db	722	CGCTGGTACCCGCCGCTCTCATGGTGAAGGAGATGCCG	781	2 TGACCTGCGAAAGGCCAAGGGCATTGGAGCTCTCCGGGT	61	
Qy	782	ATGGCACAGGGCTACCGTCAACGGGAGGTCAACCGGTCC	841	62 TCTGGCAGAACGGCCGGCTTACCCGGCTGACACTCCAGAGCATGGCGGCC	121	
Db	782	ATGGCACAGGGCTACCGTCAACGGGAGGTCAACCGGTCC	841	62 TCTGGCAGAACGGCATTGGAGCTCTCCGGGT	121	
Qy	842	CGGGGAGGGACTTCACACTGGCACCTCGGTCAATGGAGGTGCTCCGGTGC	901	62 TCTGGCAGAACGGCATTGGAGCTCTCCGGGT	121	
Db	842	CGGGGAGGGACTTCACACTGGCACCTCGGTCAATGGAGGTGCTCCGGTGC	901	62 TCTGGCAGAACGGCATTGGAGCTCTCCGGGT	121	
Qy	902	TGATGACCCGATCTGGCTCATGGTGGAGGTGCTCCGGTGC	901	122 GGGAAACGGCTCATGGGCTATGAAAGGGGACTGATCCGGAGCTTCTGC	181	
Db	902	TGATGACCCGATCTGGCTCATGGTGGAGGTGCTCCGGTGC	901	122 GGGAAACGGCTCATGGGCTATGAAAGGGGACTGATCCGGAGCTTCTGC	181	
Qy	962	ATGTGGATGAATTTCAAATCTAGGTGGCTCATGGTGGAGGTGCTCCGGTGC	1021	122 GAAACATCATCGGGCCGACGTGAAGGCAACCAGGT	241	
Db	962	ATGTGGATGAATTTCAAATCTAGGTGGCTCATGGTGGAGGTGCTCCGGTGC	1021	122 GAAACATCATCGGGCCGACGTGAAGGCAACCAGGT	241	
Qy	1022	TTCGGGAAGGGGCTTGGCAGTTACACACCGGATATCAGGAGTGGGGGGCC	1081	182 GAAACATCATCGGGCCGACGTGAAGGCAACCAGGT	241	
Db	1022	TTCGGGAAGGGGCTTGGCAGTTACACACCGGATATCAGGAGTGGGGGGCC	1081	182 GAAACATCATCGGGCCGACGTGAAGGCAACCAGGT	241	
Qy	1082	ACGACTCGATGTTCATCGCCCAGGTCAAGGAGCTGGCTGGGAGCTGGT	1141	182 GAAACATCATCGGGCCGACGTGAAGGCAACCAGGT	241	
Db	1082	ACGACTCGATGTTCATCGCCCAGGTCAAGGAGCTGGCTGGGAGCTGGT	1141	182 GAAACATCATCGGGCCGACGTGAAGGCAACCAGGT	241	
Qy	1142	CTGGCCGGACGCCACGGTCCCGGAAGAATGGGACGGATTGACCTTCGCAAGGAACTG	1201	182 GAAACATCATCGGGCCGACGTGAAGGCAACCAGGT	241	
Db	1142	CTGGCCGGACGCCACGGTCCCGGAAGAATGGGACGGATTGACCTTCGCAAGGAACTG	1201	182 GAAACATCATCGGGCCGACGTGAAGGCAACCAGGT	241	
Qy	1202	AGGTTGGTCCCGGGTAATGAAAGGGCACTTGACGGTAA	1239	182 GAAACATCATCGGGCCGACGTGAAGGCAACCAGGT	241	
Db	1202	AGGTTGGTCCCGGGTAATGAAAGGGCACTTGACGGTAA	1239	182 GAAACATCATCGGGCCGACGTGAAGGCAACCAGGT	241	
Qy	602	AGCAGGAGAAATAAGAGATCGGAGCTTCTGGGCTCATCGAACAGGAACT	661	182 GAAACATCATCGGGCCGACGTGAAGGCAACCAGGT	241	
Db	602	AGCAGGAGAAATAAGAGATCGGAGCTTCTGGGCTCATCGAACAGGAACT	661	182 GAAACATCATCGGGCCGACGTGAAGGCAACCAGGT	241	
US-09-916-501-6				US-09-916-501-6		
Qy	662	Sequence 6, Application US/0916501		Sequence 6, Application US/0916501		
Db	662	Patent No. US20020132848A1		Patent No. US20020132848A1		
Qy	722	GENERAL INFORMATION:		GENERAL INFORMATION:		
Db	722	APPLICANT: KRIMMER, Hans-Peter		APPLICANT: KRIMMER, Hans-Peter		
Qy	722	REICHERT, Dietmar		REICHERT, Dietmar		
Db	722	DRAUZ, Karlheinz		DRAUZ, Karlheinz		
Qy	722	KLEMENT, Ingo		KLEMENT, Ingo		
Db	722	APPLICANT: MAY, Oliver		APPLICANT: MAY, Oliver		
Qy	782	TITLE OF INVENTION: Process for the Preparation of Allysine Acetal		TITLE OF INVENTION: Process for the Preparation of Allysine Acetal		
Db	782	FILE REFERENCE: 210740US-10757-9350-0-X		FILE REFERENCE: 210740US-10757-9350-0-X		
Qy	782	CURRENT APPLICATION NUMBER: US/09/916,501		CURRENT APPLICATION NUMBER: US/09/916,501		
Db	782	PRIOR APPLICATION FILING DATE: 2001-07-30		PRIOR APPLICATION FILING DATE: 2001-07-30		
Qy	842	PRIOR APPLICATION NUMBER: Germany 100 37 115.9		PRIOR APPLICATION NUMBER: Germany 100 37 115.9		
Db	842	NUMBER OF SEQ ID NOS: 6		NUMBER OF SEQ ID NOS: 6		
Qy	902	SOFTWARE: Patentin version 3.1		SOFTWARE: Patentin version 3.1		
Db	902	SEQ ID NO: 6		SEQ ID NO: 6		
Qy	902	LENGTH: 1263		LENGTH: 1263		
Db	902	TYPE: DNA		TYPE: DNA		
Qy	902	ORGANISM: Arthrobacter aurescens		ORGANISM: Arthrobacter aurescens		
Db	902	FEATURE: misc_feature		FEATURE: misc_feature		
Qy	962	NAME/KEY: misc_feature		NAME/KEY: misc_feature		

Db	962	ATGGGATGAATTTCAACTCAGCCCCGGCTCGAGTTACACACCCGGATATCAGCAGTGGGACGCCG	1021	Qy	301	GCCCTTGAGGGCTGCCGGGTGATGCTGGAGAACGGCTACGTGAATGGCATCCATTGAG	360
Db	1022	TTCGGAAAGGGCTCGGCCCTGCAGTTACACACCCGGATATCAGCAGTGGGACGCCG	1081	Qy	301	GGCTCGAGGGCTGCAGGGATATGTGAACCGTCACTCTCTCGAG	360
Db	1022	TTCGGAAAGGGCTCGGCCCTGCAGTTACGGCACCCGGATATCAGCAGTGGGACGCCG	1081	Qy	361	TTCATCGGATCGTGGCATGGCACTGGGCGCTTCAGCAGTGGCATGGGGCGGG	420
Db	1082	ACGACTCGATGTTCATGCCCAGGTCAGGACGTGGAAATGGTTTCTGTTCAAGCCGTG	1141	Db	361	GTCACTCGGATCGTGGCATGGGGCGCTTCAGCAGGGATGGGGCTCGAG	420
Qy	1082	ACGACTCGATGTTCATGCCCAGGTCAGGACGTGGAAATGGTTTCTGTTCAAGCCGTG	1141	Qy	421	GCATTGCAAGGGTGGTGGCATGGGACTCTTGGTGTGAGGATGGAGTG	480
Db	1082	ACGACTCGATGTTCATGCCCAAGGTCACTGGACGTGGAAACTGGGAAACTG	1201	Db	421	GGATCGGGCTCGGAGACCTGGACTCTTGGTGTGAGGAGGGCTGG	480
Qy	1142	CTGGCGGAGCCACGTTCCCGAAGAAATGGACCGATTCGATGACCTTCGAAAGGAACCTG	1201	Qy	481	TCCGTTAGGCAGGGGGACTGCTTCGGCTTGAGGCCGGGAAACTGGAGGTGAG	540
Db	1142	CTGGCGGAGCCACGTTCCCGAAGAAATGGACCGATTCGATGACCTTCGAAAGGAACCTG	1201	Db	481	ACGGTGGGGCCACGGCCTTCGGGTGAAACCGGGTGAACGGGGCC	540
Qy	1202	AGGTTGTCCTCCGGTAATGAAAGGGACTTGACCGG	1236	Qy	541	CGCTCCGGGGACCTGGCTTATGAAACTACACATTGAAACAGGACCGATCCTC	600
Db	1202	AGGTTGTCCTCCGGTAATGAAAGGGACTTGACCGG	1236	Db	541	CATGAGTGGCACATGGAGGGATGACCTTCATGGCTTCAACGAGTGGGATCCTC	600
Qy	601	GAGCAGGGCAAATAAGAGATGGAGTTGTAACCTCCATCGTTGGCTTCGGCATTGGGG	660	Qy	601	GAGCAGGGAAAGGGAGATCGGGCTGATCGGGGATTCGGCTTCGGG	660
Db	601	GAGCAGGGAAAGGGAGATCGGGCTGATCGGGGATTCGGCTTCGGG	660	Db	601	GTTGCCGTCAAAGGCCAGAACCCCCATGACCTGACCTGGCCAGGAT	720
Qy	661	ATCACGGTGGAGGGAGGGGACGACGGGACGACCCCCATGACCTGGCAGGAC	720	Qy	661	ATCACGGTGGAGGGAGGGGACGACGGGACGACCCCCATGACCTGGCAGGAC	720
Db	661	ATCACGGTGGAGGGAGGGGACGACGGGACGACCCCCATGACCTGGCAGGAC	720	Db	721	CGCTGGTACCGCCGCTCATGGTGGAGGTCAACGGATTCGCC	780
Qy	721	CGCTGGTACCGCCGCTCATGGTGGAGGTCAACGGATTCGCC	780	Qy	721	CGCTGGTACACTGGACCTGGCTCCGGCATGAGGAGTCGCTCCGGGTG	900
Db	721	CGCTGGTACACTGGACCTGGCTCCGGCATGAGGAGTCGCTCCGGGTG	900	Db	841	CCGGGGAGGGTGGACTTCACACTGGACCTGGCTCCGGGTG	900
Qy	781	GATGGCACAGTGGCTACCGTTGGCCACCTCACAGTGGCCACCAACAGGTC	840	Qy	901	CTGATGCCGCGCATCTCGGTCAATCTCAGGCCATGAGGAGTCGCTCCGGGTG	960
Db	781	GACGGCACGGTGGCAACCTCACGGTCAACCTGGCTCAACCGGT	840	Db	901	CTGATGCCGCGCATCTCGGTCAACGGGATGGTGGCAAGTGGCTCGAGG	960
Qy	901	CTGATGCCGCGCATCTCGGTCAATCTCAGGCCATGAGGAGTCGCTCCGGGTG	960	Qy	961	GATGGGATGAATTTCATGGTGGCTCCAGGGCTGGCTCCGGGTG	1020
Db	901	CTGATGCCGCGCATCTCGGTCAACGGGATGGTGGCTGGGATTCGCG	960	Db	961	CGGGTGAACGGGTTCTCGGCTCAGGCCCTGGGCTCTGGGATTCGCG	1020
Qy	1021	GTTCGGAAGGGCCTGGCTTACACCGGGATATCAGCAGTGGGGGG	1080	Qy	1021	GTTCGGAAGGGCCTGGCTTACACCGGGATATCAGCAGTGGGGGG	1080
Db	1021	GTTCGGAAGGGCCTGGCTTACACCGGGATATCAGCAGTGGGGGG	1080	Db	1081	CACGACTCGATGGTCATGCCGAGCTGGCTTCAACCCATGGGACAT	1140
Qy	1081	CACGACTCGATGGTCATGCCGAGCTGGCTTCAACCCATGGGACAT	1140	Db	1081	CACGACTCGATGGTCATGCCGAGCTGGCTTCAACCCATGGGACAT	1140
Db	1	GTGACCTGCGAGAACGGCAAGGGCGCATTGAGAAAAGAGATCCGGGAGCTTCGGGG	60	Qy	1141	GCTGGCGGAGCCACGTTCCGAGAAATGGACCGATTTCGATGACCTTCGAAAGGAAC	1200
Qy	1	GTGACCTGCGAGAACGGCAAGGGCGCATTGAGAAAAGAGATCCGGGAGCTTCGGGG	60	Db	1141	GCTGGCGGAGCCACGTTCCGAGAAATGGACCGATTTCGATGACCTTCGAAAGGAAC	1200
Db	61	TTCTCGGAGAACGGCGGGCGATCGGAGGGCTCTGGACTCTCCGG	120	Qy	1201	GAGGTTGTCCTCCGGGTAATGAAAGGGACTTGACCGGTAA	1239
Qy	61	TTCTCGGAGAACGGCGGGCGATCGGAGGGCTCTGGACTCTCCGG	120	Db	1201	GATGGGGCATGGCCATTGACGGCTGACGGGGCTTCAGGTCG	1239
Db	121	CGGAAACATCATGGGGCTATGAAAGGGCCTTGAGGCTGAGGACGCC	180	Qy	181	GGAAACATCATGGGGCACTGGGAGCTGAGGCTTCTGGCATGGGGTGG	240
Qy	181	GGAAACATCATGGGGCACTGGGAGCTGAGGCTTCTGGCATGGGGTGG	240	Db	181	GGCAACATCATGGGGCACTGGGAGCTGAGGCTTCTGGCATGGGGTGG	240
Qy	241	TCACACTTCGATTCGTCGAAACGGGGGATGTTGATGGCATGGGTG	300	Qy	241	TCACACTTCGATTCGTCGAAACGGGGGATGTTGATGGCATGGGTG	300
Db	241	TCGCACTTCGACTGGTCCGAAACGGGGGATGTTGACGGGGCTTC	300	Db	241	TCGCACTTCGACTGGTCCGAAACGGGGGATGTTGACGGGGCTTC	300

RESULT 5
US-10-289-360-3
; Sequence 3, Application US/10289360
; Publication No. US20030109013A1
; GENERAL INFORMATION:
; APPLICANT: TAKENAKA, YASUHIRO
; APPLICANT: SUZUKI, SHUNICHI
; APPLICANT: ONISHI, NORIMASA

APPLICANT: YOKOZEKI, KENZO
 TITLE OF INVENTION: DNA ENCODING HYDANTOINASE, DNA ENCODING N-CARBA-MYL-L-AMINO ACID
 TITLE OF INVENTION: HYDROLASE, RECOMBINANT DNA, TRANSFORMED CELL, METHOD OF PRODUCING
 TITLE OF INVENTION: METHOD OF PRODUCING OPTICALLY ACTIVE AMINO ACID
 FILE REFERENCE: 229195US0CONT
 CURRENT APPLICATION NUMBER: US/10/289,360
 PRIOR APPLICATION NUMBER: PCT / JP02/02173
 PRIOR FILING DATE: 2002-03-08
 PRIOR APPLICATION NUMBER: JP 2001-0650814
 PRIOR FILING DATE: 2001-03-08
 PRIOR APPLICATION NUMBER: JP 2001-298619
 PRIOR FILING DATE: 2001-09-27
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 3
 LENGTH: 1239
 TYPE: DNA
 ORGANISM: Microbacterium liquefaciens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1) .. (1239)
 OTHER INFORMATION:
 S-10-289-360-3

Query Match	Score	DB	Length
Best Local Similarity	739.8	15	1239;
Matches 927; . Conservative	0; Mismatches 312; Indels 0; Gaps 0;		
1 GTGACCCCTGCAGAAAGCGCAAGGGCGCATTGAGAAAGAGATCGGGAGCTCTCCGG 60	Qy	1081	CACGACTCGATGTTCATGCCCAAGGTACGGACGTGGATGGTTTCTGTTCCAAGCCGT 1140
1 GTGACGGCTGCAGCAGGGCCGGCGATCGGATCATGGAGGAGCTCTGGACTCTCCGG 60	Db	1081	CACGACTCGATGTTCATGCCCAAGATCACGGACGTGGATGGTTTCTGTTCCAAGCCGT 1140
61 TTCTCGGCAGAAGGCCCGGGTGTACCCGGCTGACTACACTCCAGAGCATGCCGG 120	Qy	1141	GCTGGCGAGGCCACGTTCCCGAAGAATGGACCGATTCTGATGACCTTCTGAAAGGA 1200
61 TTCTCGGTCAAAGGGGCCGGGACACGGTCTCACGTACACTCCGGACGGAGCT 120	Db	1141	GCCCCGAAAGCCATGTGCCGGAGGAATGGTCCGATTGACGATCTGGGAAGGGAGC 1200
121 CGGGAAACGGCTCATTCGGCTATGAAAGGGCCCTTGAGGGTTCTGTAAGGGCACTC 180	Qy	1201	GAGGTGTCTCCGGGTTAATGAAAGGGCACTTGACCCGGTAA 1239
121 CGAGGGTGTATCGTGGCCATGCAAGGGCATGGTCCACAGGGACGGCTCTC 180	Db	1201	GATGTGTCTTCACGTCTGACGGGCTTGACCCGGTGA 1239
181 GGAAACATCATGGCCGACGTGAAGGGCACTGATCCGGAGCTTCTGGATCGGGTGGT 240	Qy	RESULT 6	
181 GGCAACATCATGGTGGCTGAGGGAGGGACCCGGCTCTGGCGATGCCCTCGGC 240	Db	US-09-950-772-7	
241 TCACACTTGGATTCTGTCGACTCGGTCCGAACGGGGATGTTGATGGCACTGCAGGGCTGGTGTGC 300	Qy	Sequence 7, Application US/09950772	
241 TCGCACTTGGATTCTGTCGACTCGGTCCGAACGGGGATGTTGATGGCACTGCAGGGCTGGTGTGC 300	Db	Patent No. US2002102713A1	
301 GCCCTTGAGGCTCGGGGATGCTGGAAACGGGCTACGTGAATGGCATCCATTGAG 360	Qy	GENERAL INFORMATION:	
301 GCGCTCGAGGCTCGGGGATGCTGGAAACGGGGATGTTGAAACCGGACGGCATCCCTCGAG 360	Db	APPLICANT: SUZUKI, Shunichi	
361 TTCATCGCGATCGTGGAGGGAAAGGGGATGCTGGAAACGGGCTACGTGAATGGCATCCATTGAG 360	Qy	APPLICANT: ONISHI, No. US20020102713Alimasa	
361 GTCATCGCGATCGTGGAGGGCAACGGGGATGTTGAAACCGGACGGCATCCCTCGAG 360	Db	APPLICANT: YOKOZEKI, Kenzo	
421 GCCATTGCAGGGTTGGTCGCCGACAGGGAAACTGGACTCTTGTGATGGGATGGAGTG 480	Qy	TITLE OF INVENTION: 5-SUBSTITUTED HYDANTOIN RACEMASE, DNA CODING FOR THE RACEMASE, PROCESSES FOR PRODUCING OPTICALLY ACTIVE AMINO ACIDS	
421 GCGATCGGGCTCGTGTGGACCGGGCATCTGGACAGGGAGAAGACGGCTG 480	Db	FILE REFERENCE: 212318US0	
481 TCCGTTAGGCAGGGGGCTACTGCCTTCGGCTTATGAACTACACATTGAAAGGACGGCTGCC 540	Qy	CURRENT APPLICATION NUMBER: US/09/950,772	
481 ACGGTGCGGAGGGCCACGGGCTTCGGCTGAACGGGGTGAAGACGGCTGCC 540	Db	PRIOR APPLICATION NUMBER: JP2000-09-13	
541 CGCTCCGGGGGACCTCGGGACTCGGGTTATGAACTACACATTGAAAGGACGGCTGCC 600	Qy	PRIOR APPLICATION NUMBER: JP2000-278571	
541 CGTACGAGGGATGACCTTCGGGCTTCATCGAGTGCACATCGAGGGGGCATTCCTC 600	Db	PRIOR APPLICATION NUMBER: JP2001-65815	
601 GAGCAGGAGCAAATAGAGATCGGAGTTGAAACCTCCATCGTGGCGCATGGGG 660	Qy	PRIOR FILING DATE: 2001-03-08	
601 GAGCAGGAGAAAGGTGGAGATCGGGTGTGACGGGGATCTGGACTCTCCCGC 2164	Db	NUMBER OF SEQ ID NOS: 12	
Query Match Score 739.8; DB 9; Length 3343;		SOFTWARE: PatentIn version 3.1	
Best Local Similarity 74.8%; Pred. No. 1e-211;		SEQ ID NO 7	
Matches 927; Conservative 0; Mismatches 312; Indels 0; Gaps 0;		LENGTH: 3343	
ORGANISM: Microbacterium liquefaciens		TYPE: DNA	
US-09-950-772-7		US-09-950-772-7	

QY	61	TTCCTGGCAGAAGGCCCGTGTACCCGGTCAACTCCAGAGGCAAGCTACACTCGCAGGCCCGG	120	QY	1141	GCTGGCCGCCACCTTCCCGAAGAAATGGACCGATTTCGATGACCTTCGATGACTTCGAAAGGAAC	1200
DB	2165	TTCTGGTCGAAGGCCCGGTGACACTGTCTCACGTACACTCCGGAGCACGCCGCCG	2224	DB	3245	GCGGGGGAAGCCATGTCGGGAAATGGCTCGATCTGGGAGAATGGCTCGATTTCGACGATCTGCGGAAGGGGAC	3304
QY	121	CGGAAACGCTCATTGCGGCTATGAAAGGGCCCGCTTGAGCCTTCGTTGAGAACGCACTC	180	QY	1201	GAGGTGGTCTCCGGGTAATGAAGGGCACTTGACCGGTA	1239
DB	2225	CGAGAGGTGATGGCCCATGAGGGCATGGGACGGGCTGAGGACGGGACGGGCTC	2284	DB	3305	GATGTGGTCTTACAGTCGTCGACGGGCTGACGGCTGACCGGCTGA	3343
QY	181	GGAAACATCATGGCCGACGTGAAGGCCACTGATCCGGAGCTTCTCGGATCGGGTCGGT	240	RESULT 7			
DB	2285	GGCAACATCATGGTCGGGTGAGGGGACCCGGTCTGCCTTCGGC	2344	US-10-289-360-5			
QY	241	TCACACTTCGATTCTGTCGAACGGGGATGTTGATGGCACTGAGGGTGGTGC	300	; Sequence 5, Application US/10289360			
DB	2345	TGGCACTTCGACTCGGTGAGGGTCTGCAGGGCACCCGGGATGTTGAGGGTGGTGC	2404	; Publication No. US20030109013A1			
QY	301	GCCCTTGAGGCTGCCCGGTGATGCTGGAGAACGGTACGTGAGATCGGCACTCCATTGAG	360	; GENERAL INFORMATION:			
DB	2405	GCGCTCGAGGCTGCCGGAGGGTCTGCAGGGAGGGATGTTGAGGGTATGTGAAACGGTCACTCTCTCGAG	2464	; APPLICANT: TAKENAKA, YASUHIRO			
QY	361	TTCATCGCGATGTTGGAGAACGGGGCTTACAGGTGGCATGTTGGGGCGG	420	; APPLICANT: SUZUKI, SHUNICHI			
DB	2465	GTCATCGCGATGTTGGAGGGCACCCGGTTACAGGTGGGATGGAGTG	2524	; APPLICANT: ONISHI, NORIMASA			
QY	421	GCCATTGCAAGGGTGTGGTCAAGGGAAACTGGGACTCTTTGGTTGATGAGGTGGAGTG	480	; APPLICANT: YOKOZEKI, KENZO			
DB	2525	GCGATCGGGGGCTGTCGAGGGGGCACCCCTGGACACCGCAGGGCTG	2584	; TITLE OF INVENTION: DNA ENCODING HYDANTOINASE, DNA ENCODING N-CARBAMYL-L-AMINO ACID			
QY	481	TCCGTTAGGCAGGGGTACTRGCCRTGCGCTGAAGGGGGAAACTGCAAGGGTGCAGCC	540	; METHOD OF PRODUCING OPTICALLY ACTIVE AMINO ACID			
DB	2585	ACGGGGGCCACGGGCCACGGGCTTGGCTGAAACGGGTGAGCTGGGACGGGCC	2644	; FILE REFERENCE: 229195US0CONT			
QY	541	CGTCCGGGGGACCTGCGTGTGCTTATGAAACTACACATTGAAACAAGGACCGATCC	600	; CURRENT APPLICATION NUMBER: US/10/289,360			
DB	2645	CGTAGAGGGATGACCTGGCCCTTCATCGAGTGGCACATCGAGGAGGGCGATCC	2704	; CURRENT FILING DATE: 2002-11-07			
QY	601	GAGCAGGCAAATAGAGATCGGAGTTGTAACCTCCATCGTGGCATTTGGG	660	; PCT/JP02/02173			
DB	2705	GAGCAGGAAAGGTGGGATCGGGGATCGTGGTGTGGCTGGCTTCCGG	2764	; PRIOR APPLICATION NUMBER: JP 2001-0650814			
QY	661	GTTGCCGTCAAAGCGAACCGGCCAACCCCCATGCACTTGGCAGGAT	720	; PRIOR FILING DATE: 2001-03-08			
DB	2765	ATCACGGTGGAGGGCAGGACCGGCCAACGGGGACGCCCCATGCACTGGCAGGAC	2824	; PRIOR APPLICATION NUMBER: JP 2001-09-27			
QY	721	GCGCTGCTACCCGGCTCTCATGGTGGGACCTCACAGTGGAGGTAAACGGAGATCGCC	780	; NUMBER OF SEQ ID NOS: 14			
DB	2825	GCGCTGCTGGGGCTCATGGTGGGCTCATGGTGGAGATCAATGGTCAACGGAGATCGCG	2884	; SOFTWARE: PatentIn version 3.1			
QY	781	GATGGACAGTGGCTACCGTTGGCACCTCACAGTGGAGGTAAACGGAGATCGCC	840	; LENGTH: 3343			
DB	2885	GACGGCACGGTGGGACCGTGGCCACCTCACGGTGGCTCAACGGAGTCAACGGGTT	2944	; SEQ ID NO 5			
QY	841	CGGGGGAGGGTGGACTTCACACTGGACCTCGGTTCGGCATGAGGAAGTCGCTGGT	900	; TYPE: DNA			
DB	2945	CCGGGGGGCTGTCAGGTTACGGCTCGATCTGGCATGCCAACGGAGGACTCGATCCGG	3004	; ORGANISM: Microbacterium liquefaciens			
QY	901	CTGATCGACCCGATCTCGGTCACTGGCTGGGAGGGTGGCTGCC	960	Query Match 59.7%	Score 739.8;	DB 15;	Length 3343;
DB	3005	CGGTGACAGGGATCGAGGGCATGGTGGCAAGGGGGCTGTTGAGGCACTTCGGTGTG	1020	Best Local Similarity 74.8%;	Pred. No. 1e-211;		
QY	961	GATGTGGATGAATTTCATACTCGCCGGCTTGGCAGTTCGCTACCATGGGAGGCC	1020	Matches 927; Conservative 0;	Mismatches 312;	Indels 0;	Gaps 0;
DB	3065	CGGTGAACGGGTCTTGGCAGGGCATGGTGGCAAGGGGGCTGTTGAGGCACTTCGGTGTG	3064	QY 1	GTGACCCCTGGAGAAAGGCCAACGGGGCAATTGAGAAAGAGATC	GTGGCTCTCCCGG	60
QY	1021	GTTCGGAAAGGGCCCTGGCAGTTCGCTGGGAGGGATATCAGCAGGGGGATTC	1080	DB 2105	GTGACGCTGAGGGCTGAGGGGGCTGATCGGACTCTGGGACTCTCCCG	2164	
DB	3125	GTGGGGAAGGGCTTGGGAAACTGGCTCATGGCCAACTGGGAGGGGGCTG	3124	QY 61	TCTCGGGAGGGCCCTGGCTGACCTCCAGAGCATGGGGCCGGCG	120	
QY	1081	CACGACTCGATGTTCATGCCAGGTGGCAAGCTGGGAAATGGTTTCAGGCAAGCCG	1140	DB 2165	TCTCGGGTGAAGGGGGCTGACCGTCACTCCGGACCGGAC	2224	
DB	3185	CACGACTCGATGTTCATGCCAGATCCAGGCTGGGACACCCGGCTG	3244	QY 121	CGGAAACATCATCGGCCGACGTGAAGGGACTGATGGGGCTATGGGAGGCACTC	180	
QY	1141	GCTGGCCGCCACCTTCCCGAAGAAATGGGACTCTGGGACTGAGGGCTGGT	300	DB 2225	CGAGAGGTGATCGTCGGCCATGCAAGGGGGCTGAGGGCTCACGGCTCTCGAGGCTC	2284	
DB	3245	GCGGGGAAGGGCCACCCGGCTTGGGAGGGGGCTGAGGGGGCTG	3064	QY 181	GGAAACATCATCGGCCGACGTGAAGGGACTGATGGGGCTATGGGAGGCACTC	240	
QY	1201	GAGGTGGTCTCCGGGTAATGAAGGGCACTTGACCGGTA	1239	DB 2285	GGAAACATCATCGGCCGACCCGGCTGAGGGGGCTGAGGGGGCTGAGGGGGCTG	2344	
DB	3305	GATGTGGTCTTACAGTCGTCGACGGGCTGACGGGCTGACCGGCTG	3124	QY 241	TCAACACTTGGATTCTGGCCGGGTGATGCTGGACTGAGGGCTGGTGTG	300	
QY	1201	GAGGTGGTCTCCGGGTAATGAAGGGCACTTGACCGGTA	1239	DB 2345	TGGCACTTCGACTCGGGCTGGGAGGGGGCTGAGGGGGCTGAGGGGGCTG	2404	
DB	3305	GATGTGGTCTTACAGTCGTCGACGGGCTGACGGGCTGACCGGCTG	3124	QY 301	TCCACTTGGGCTGGGAGGGGGCTGAGGGGGCTGAGGGGGCTGAGGGGGCTG	360	
QY	1201	GAGGTGGTCTCCGGGTAATGAAGGGCACTTGACCGGTA	1239	DB 2405	CGGCTCGAGGGCTGGGAGGGGGCTGAGGGGGCTGAGGGGGCTGAGGGGGCTG	2464	
DB	3305	GATGTGGTCTTACAGTCGTCGACGGGCTGACGGGCTGACCGGCTG	3124	QY 361	TTCATCGGATCTGGGAGGGGGCTGAGGGGGCTGAGGGGGCTGAGGGGGCTG	420	
QY	1201	GAGGTGGTCTCCGGGTAATGAAGGGCACTTGACCGGTA	1239	DB 2465	GTCATCGGATCTGGGAGGGGGCTGAGGGGGCTGAGGGGGCTGAGGGGGCTG	2524	
DB	3305	GATGTGGTCTTACAGTCGTCGACGGGCTGACGGGCTGACCGGCTG	3124	QY 421	GGCATTTGCAAGGGAAACTGGGACTCTGGGCTGGGAGGGGGCTGAGGGGGCTG	480	
QY	1201	GAGGTGGTCTCCGGGTAATGAAGGGCACTTGACCGGTA	1239	DB 2525	GGGATCGGGGGCTCGTCCGAGACCCGGGATCTGGGAGGGGGCTGAGGGGGCTG	2584	

Qy	481	TCCGRTAGGCAGGGCTACTGCCTCGGTTGAAGGCCGGAACTGCAGGCTGCAGCC	540	; TYPE: DNA ; ORGANISM: Listeria innocua
Db	2585	ACGGTGGCGAGGGCCACGGCTTCGGGTGAAACCGGGTGAGCTGCGGAGGCC	2644	; FEATURE: ; NAME/KEY: ; LOCATION: (1)..(end) ; OTHER INFORMATION: n can be any nucleotide: a, g, c or t/u ; US-10-398-221-10
Qy	541	CGCTCAGGGGACCTGCGCTTTATCGAACTACACATTGAAACAAGGACCGATCTC	600	
Db	2645	CGTACCGGGATGACCTTCGCGCTTCATCGAGTGCACATCGAGCAGGGCGATCTC	2704	
Qy	601	GAGCAGGAGCAAATAGAGATCGGAGTTAACCTCATCGTTGGGTTCCGCATTGGGG	660	Query Match 17.8%; Score 220.6; DB 16; Length 1163020;
Db	2705	GAGCAGGAGAAAGTGGAGATGGCTCGTGCCTGACGGGATCGTGCCTTCGG	2764	Best Local Similarity 50.2%; Pred. No. 3.4e-55; Matches 572; Conservative 0; Mismatches 564; Indels 3; Gaps 1;
Qy	661	GTTGCCGTCAAGCCAGAACGGGACCCGGCAACCCCCATGCACCTGGCCAGGAT	720	Qy 73 GGCCCCGGTGTACCGGGTGAACCTCACACTCCAGAGCATGCCCGGGAAACGCTC 132
Db	2765	ATCACGGTGGAGGGCAGGGGACCCGGGAGACCCCCATGCACTGGGGAGAC	2824	Db 478591 GGTCAAGGGAAACAACCTGGCTTACATATAGCAAAGAAGACCTCGGGCGGTATTATTAA 478650
Qy	721	GCGCTGGTACCCGGCGTCATGGTGAGGGAGGTCAACCGGATCGGCC	780	Qy 133 ATTGGCGCTATGAAAGGGGGCCCTTGAGCGGTTCTGTGAAAGAGCGCACTCGGAACATCATC 192
Db	2825	GCGCTGGCGGGCGTCATGGTGGAGAGTCAATCGGTTCTGTCAACCGAGATCGCG	2884	Db 478651 AAAGAAGAAATGGCTAAAGTAGGCGCTTACTGTCAGAAAGATGGATATCTAC 478710
Qy	781	GATGGCACAGTGGTACCGTGGCCACCTCACAGTGGCCCCGGTGGAGGCAACCAGGTC	840	Qy 193 GGCGGACGTGAAGGCACTGATCCGGAGCTTCTTGCCTGGATTCACACTTCGAT 252
Db	2885	GACGGCACGGTGGGACCGTGGCCACCTCACGGTGACCCCCTGGGCTCAACCAGTT	2944	Db 478711 GGACGACTAGAAGGGACAATCAGACATAACCGAGTAATCGTGGGTATCTCATTTGAC 478770
Qy	841	CGGGGGAGGGGAACTTCAACACTGGACCTTGCGTTCTCCGGCATGGAGTGCCTCCGGCTG	900	Qy 253 TCTGTCCGAAACGGGGGATGTTGATGGCACTGCACTGCAGGGCTGGCTTGGGGCT 312
Db	2945	CCGGGGGGCGTCAGGTTCAACGGGATCGCTCGATCGGATCGGCCAGAGGATCGATCGCTCGATCGCTCGATCGGATCGGCC	3004	Db 478771 TCTGTGCGCAAATGGTGGCTTTCATGAACAGCAAAATAAACCCCATTTCACACTGAATCATCGCAATG 478830
Qy	901	CTGATCGACCGCATCTCGGTATGGTCCGGGAGGTGCCTCCAGGCCGGTGGCTGCC	960	Qy 313 GCCGGGGTGTATGCTGAGAAGGGCTACGTGAATGGCATCCATTGAGTTCACTCGCGATC 372
Db	3005	CTGGTGGACAGGATCGAGGGATGGGAGTGGTGGAGAAGTGCCTGGGGGCC	3064	Db 478831 GCTAGCGGTTTCATGAACAGCAAAATAAACCCCATTTCACACTGAATCATCGCAATG 478890
Qy	961	GATGTGGATGAATTTCATCTCAGCCGGCTGGCAGCTGGCTTCTACCATGGTGGAGGCC	1020	Qy 373 GTGGAGGAGGAAGGGGCCGCTTCAGCAGTGGCATGTTGGGGCATTTGAGCTGGAGGCC 432
Db	3065	GCGGTTGAACGGGTCTTCGGCTCAGGCCCTGGCTCAGGCCATCGGGTGGATCGC	3124	Db 478891 GTGGAAAGAAAGGGATCTCGTTGGAGGGACTACTTGCCTCCGAACATTACAGGA 478950
Qy	1021	GTTCGCGAAGGGCCTCGGCCCTGGCAGTTACACACCGGATATCAGCAGTGGGGGCC	1080	Qy 433 TTGGCTGGGACAGGGAAACTGGACTCTGGTGTATGGAGGTGTCGGCTTAGGCAG 492
Db	3185	CACGACTCGATGTCTACGCCATCGGCCAGATCACGGACGGGGAGGG 3184		Db 479011 GCTATGGCAAACCTAGGATTGACGCAAATCAAGTCATACGCATTGGTCCAAGAT 479070
Qy	1081	CACGACTCGATGTCTACGCCATCGGCCCTGGCAGTTACACACCGGATATCAGCAGTGGGGGCC	1140	Qy 493 GGGGCTACTGCCTTCGGCTGAAGCCGGGAACCTGCAAGGCTCCGGGGCG 552
Db	3185	CACGACTCGATGTCTACGCCATCGGCCAGATCACGGACGGGGAGGG 3184		Db 479011 GCTATGGCAAACCTAGGATTGACGCAAATCAAGTCATACGCATTGGTCCAAGAT 479070
Qy	1141	GCTGGCGAGGGCACGGTCCCGAAGAAATGGACCGATTTCGATGACCTTCGCAAAGGAACCTGGGTTCTCGTCCCAGCCAGC	1200	Qy 553 GACCTGGCTGGCTTATCGAAACTACATCGAACAGGACCTCTCAGGAGGGCAA 612
Db	3245	GCGGGGGCGAAGGCCATGTGCCGGAGGAATGGTCCGATTCGAGGAACTGGGAAGGGGAGC	3343	Db 479071 TCCGTCAAAGCTTTCAATTGAAACTACATCGAACAAAGGCCAGTGGAAATGGGAAT 479130
Qy	1201	GAGGTGTTCTCGGGTAATGAGGGTAACTGGGACTTGTGACCGTAA	1239	Qy 613 ATAGAGATCGGAGGTTGTAACCTCCATCGTTGGCTTCGGCATGGGGTTGCCGTCAA 672
Db	3305	GATGTGGTCCTCACGGTGGCTTACGTGGCCGGAGGAATGGTCCGATTCGAGGATCTGGGAAGGGTGA	3343	Db 479131 GAAGATGTTGGCTAGGTGACACGGTGGCTTAAACAGAAAATGGTAAAGTACCGTAAAG 479190
Qy				Qy 673 GGCAGAAGGCACCGCCGGCAACCCCATGACCGATGCCAGGATGGCTGGTACCC 732
Db				Db 479191 GGTCAAGGCAAGCTGGCTCATGGTGGGGAGGTCAACGGATGGCTGGTACCC 732
Qy				Qy 733 GCGGCTCTCATGGTGGGGAGGTCAACGGGTTCTGTCAACGAGATGGCTGGTACCC 732
Db				Qy 793 GCTACCGGTTGGCCACCTCACAGTGGCCCTGGAGGCAACCCAGGTCCGGGGAGGTG 792
Qy				Db 479251 GCTGTGCTCAATTCTTAAGCTTCTGTAAACTAGCTATGGCTTAAGTGC 479310
Qy				Qy 793 GCTACCGGTTGGCCACCTCACAGTGGCCCTGGAGGCAACCCAGGTCCGGGGAGGTG 852
Db				Db 479311 TTACAAATGGCAACCTTAATGCTACCCGGAAACGGGCTTAATGTAATACAAACAAAGTC 479370
Qy				Qy 853 GACCTCACACTGGACCTGGCTTCCTCCGGCAATGAGGATGTCGTGCTGACCCGC 912
Db				Db 479371 GTTTCAACGGTGGATATTGAGCAAATACATAGAAAAA 479430
Qy				Qy 913 ATCTCGGTCTAGGGTGGGAGGGCTCCAGGGCCGGTGTGGCTGGCATGGTGA 972
Db				Db 479431 A---CAAAAAAAAGTTTCAAGGCTTTCGAAAATAGAAGAT 479487
Qy				Qy 973 TTTTCAATCTAGCCGGTGGCTTACCATGGTGGACGCCGGTTGGCGAAGGC 1032

RESULT 8

US-10-398-221-10
Sequence 10, Application US/10398221
Publication No. US20040018514A1

GENERAL INFORMATION:

APPLICANT: KUNST, Frederik

APPLICANT: GLASER, Philippe

TITLE OF INVENTION: Listeria innocua, genome and applications

FILE REFERENCE: 344 702 - US

CURRENT APPLICATION NUMBER: US/10/398, 221

CURRENT FILING DATE: 2003-03-27

PRIOR APPLICATION NUMBER: PCT/FR 01/03 061

PRIOR FILING DATE: 2001-10-04

PRIOR APPLICATION NUMBER: FR 00/12 697

NUMBER OF SEQ ID NOS: 4025

SOFTWARE: PatentIn version 3.0

SEQ ID NO 10

LENGTH: 1163020

Query Match 12.8%; Score 158.6; DB 16; Length 3880;
 Best Local Similarity 49.8%; Pred. No. 4e-37;
 Matches 429; Conservative 0; Mismatches 429; Indels 3; Gaps 1;

Qy 351 TCCATTGAGTTCAATCGCGATCTGGAGGAAGGGGCCGCTTCAGCAGTGGCATGTT 410
 Db 3855 TCCGGTTGAAATTATTGCGATGGTGAAGAAAGGGCTGGCTGGCTCCCT 3796

Qy 411 GCGCGCCGGCCATTGAGGGTTGGCGACAGGGAACTGGACTCTTGGTGTGTA 470
 Db 3795 TGCTTACGTAATACTGGCAAAGTACCAAAAGAAATGCTCCATGAAAGATAT 3736

Qy 471 GGATGGAGTGTCCCGTTAGGCAAGGGGTACTGCCTCGGCTGAAGCGGGCGAAC TGCA 530
 Db 3735 AGATGGAATTACAGGCCAGAAGCCATGGCAAAGTAGGATTGATGCCAAATCAAGTAGT 3676

Qy 531 GGCTGAGCCCGTCCGGGACCTGGCTTATCGAACTACACATTGAAACAAGG 590
 Db 3675 TACAGCTATCCGACGAAAAGAATCCGTTAACGATTTATCGAACTACACATTGAGCAAGG 3616

Qy 591 ACCGATCCTCGAGCAGGCAAATAGAGATCGGAGTTGAACCTCCATGGTGGCGTTCG 650
 Db 3615 ACCAGTCCCTAGAAAATGCCATTAGAAGATGTGGCGCTAGTAGATACTAGTAGTCGGTTAAC '3556

Qy 651 CGCATGGGTTGCCGTCAAAGGCAGAACGGGACAACCCCCATGCACCT 710
 Db 3555 AGAAATAAAAGTAACAGTAAAGACAAGGGCAAAAGGGCCATGGCAGGCAAACCCCAATGCTGA 3496

Qy 711 GGGCAGGATGCCGATGGCTGGTACCCGGCTCTCATGGTGAGGGAGGTCAACCGGGTTCGTCAA 770
 Db 3495 CCGAAAAGATGCCGTGATCACAGTGGTCAAGGAAATTAGGTCAACTACCAGAACTAGGTAT 3436

Qy 771 CGAGATGCCGATGGCACAGGGCTACCGGCTCACAGTGGCCACCTCACAGTGGCTTGGGAGG 830
 Db 3435 CCAAGAAAGGGGGAACTGTATTAAACGTGGCAAAACTCAACGTCTATTCAAATGGGC 3376

Qy 831 CAACCAAGTCCGGGGAGGGGACTTCAACTGGACCTGGCTCGGTACGGAGGAGTC 890
 Db 3375 AACAGTTATTCCAGATAAAAGTCGTTTACCGTAGATATCCGAGAAAAGACGAAATTC A 3316

Qy 891 GCTCCGGTGTGATCGACCGCATCTCGGTCAATGGTGGCTCGAGGTGGCTCCAGGCCGG 950
 Db 3315 CGTCCAAAATACATTAGCAAAAGAAATTAT--CCATCCGCAAGAAAAACGG 3259

Qy 951 TGTGGCTGCCGATGGATTTCAATTCTAGCCGGGACTGGCTACCGGACCTACCAT 1010
 Db 3258 CATCACTTGGAAATAGAAGATATGCTATACGAACCGCCGACCCATTATCAAAGAGAT 3199

Qy 1011 GTGGGACGCCGTTGCCGAAGGGCCTCGGCTTCAGTCAACACCCGGATATCAGCAG 1070
 Db 3198 TCATCAAGCATTGACTGAAAGTGCCTAACCTCGCTAAATAACGAACAAATGGTAG 3139

Qy 1071 TGGGGGGGCCACGACTCGATGGTCACTGCCAGGTCAACGGGACCTGGGAATGGTTTGT 1130
 Db 3138 CGGGCGGGACAGATGGCAATTACCGAAGTAGGTGATTTGCAAGTGGCTGATTGT 3079

Qy 1131 TCCAAGCGGTGCTGGCCGAGGCCACGTTCCCGAAGAATGGACCCGATTTCGATGACCTTCG 1190
 Db 3078 CCCTAGCCACAATGGTATAAGCCATGGCAATGCAAGTAGGATTACGACAAAACCTCCA 3019

Qy 1191 CAAAGGAACGTGAGGTGTCCT 1211
 Db 3018 AAAAGGAATCGAAAGTCGATGCTACT 2998

RESULT 11
 US-09-94-300-308
 ; Sequence 308, Application US/09974300
 ; Patent No. US20020146721A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Berka, Randy M.
 ; APPLICANT: Clausen, Ib Groth

; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
 ; TITLE OF INVENTION: Expression
 ; FILE REFERENCE: 10085.500-US
 ; CURRENT APPLICATION NUMBER: US/09/974,300
 ; CURRENT FILING DATE: 2001-10-05
 ; PRIOR APPLICATION NUMBER: 09/680,598
 ; PRIOR FILING DATE: 2000-10-06
 ; PRIOR APPLICATION NUMBER: 60/279,526
 ; PRIOR FILING DATE: 2001-03-27
 ; NUMBER OF SEQ ID NOS: 8481
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 308
 ; LENGTH: 1218
 ; TYPE: DNA
 ; ORGANISM: Bacillus licheniformis
 ; US-09-974-300-308

Query Match 12.1%; Score 150.2; DB 9; Length 1218;
 Best Local Similarity 47.3%; Pred. No. 1e-34;
 Matches 492; Conservative 0; Mismatches 543; Indels 6; Gaps 1;

Qy 180 CGAAAACATCATGGCCGACGTGAAGGCACTGATCAGGCTTCCTGAGATCGGGCTCGG 239
 Db 167 CGAAATCTGTTGGAGAATAGAAGGTACAGAGAACGTAACAGAACGAAAC 226

Qy 240 TTACACACTTCGATTCTGTCGGAAACGGGGGATGTTGATGGCACTGAGGCTGGTGTG 299
 Db 227 TTACACACATTGATAGGTCAACGGGAAATTGACGGGATGGATTAGC 286

Qy 300 CGCCCTGAGGCTGCCGGGTATGCTGGATACTGGCATCCATTGA 359
 Db 287 GAGCCTTGTGCTGCAAAACACTTCTGGACACGTAAGGCCAGTGGCATGTTGGGATCGGA 346

Qy 360 GTTCATCGCGATCGTGGTGGCCGACAGGGAACTGGACTCTTGGTGTGAGGATGGAGT 419
 Db 347 AGTCGTTGTCTCTGTGGCAAGGGAAAGGGAAAGGGGATGGGATCGGG 406

Qy 420 GGCCATTGGAGGTGGTGGCCGACAGGGAACTGGACTCTTGGTGTGAGGATGGAGT 479
 Db 407 GAATTAAACGGGCTTGTGACTGTCAAAAGACGCCGGCTGTGCAAGATCGGGAGGAT 466

Qy 480 GTCCGTTAGGCAGGGGACTGGCTACTGGCTTCGGTTATCGAACAGGGATCC 539
 Db 467 ATCGATTGAGCGGGATGATGAAAGTGGCTTATCGAGCTTATGGGATCGGG 520

Qy 540 CGCTCCGGGGAAATAGAGATGGGACTTCACTGGCTTCACTGGCTTCGGCATGGC 659
 Db 521 CCCGTTCAAGGGTGTGATGAAAGTGGCTTATCGAGCTTATGGGAGGATCT 580

Qy 600 CGAGCAGGCAAAATAGAGATGGGACTTCACTGGCTTCACTGGCTTCGGCATGGC 659
 Db 581 TGAGGAGGGGGCTCAAATGGGATTTGTCAGGACATCGTCGGACAAAGACGGTTAAC 640

Qy 660 GTGGCCGTCAAAGGAGGCAACGGGACCCAGGGGACACACCCCATTGGCAGGCCAGGA 719
 Db 641 GGTTATGGCTCAAGGGAAACCACTGGGGACGAGCCGATGAAACGCCAGAAAAGA 700

Qy 720 TCGGCTGGTACCCGGCTCTCATGGTGAGGGAGGTCAACGGGTCTGTCAAAGGATCGC 779
 Db 701 TGCGGTTGGCGGTCTGGCATTTGTCATGGCTTCAAGGGCTAAAGGCTGCTTA 760

Qy 780 CGATGGCACAGTGGTACCGTTGGCCACCTCACAGTGGCCACCCGGAGGATGAGG 839
 Db 761 TCCGGTTAAACAGCAACGGTGGGAGGATGAGGAAATGTGCGGAATGTCTCAT 820

Qy 840 CCCGGGGAGGGTGGACTTCACACTGGACCTGGCTTCGGCATGGAGGTGCGCTCCGGCGT 899
 Db 821 ATCAGGAGAAGCTGTTTCTCTGATCGGGCATCATCATGATCGGGTGGACCG 880

Qy 900 GCTGATCGACCGCATCTGGCTCATGGTGGAGGTGCGCTCCAGGGCGGTGTTGCG 959
 Db 881 GTATTGGCACGATATTGCTTCAACAGAGCTGGCCGAAATGGAAACGTTCAAGAT 940

Qy 960 CGATGGGATGAATTTCAACTCTAGCCCCGGTGCAGCTGGCTCTACCATGGGAGCGC 1019
 Db 941 CACAGTGAACAGACGACGGATGCGAACGGCTGTGAGGAATTGATTCAACT 1000
 Qy 1020 CGTTCGGAAGCGGCCCTGGCAGTTCACACACCGGATATCAGCACTGGGG 1079
 Db 1001 CTCAAAACAGTCGGCCGAGGATCTTGACGTATCCTATCAGGAGATGGTCAGGGGGGG 1060
 Qy 1080 CCACCGACTCGATGTTCATCGGCCAAGGTACGGACGTGGGAATGGTTCAAGCG 1139
 Db 1061 GCACGATGCCCAAGGTTGGCCGATGCCGACTGTGTTCCGAGCG 1120
 Qy 1140 TGCTGGGGAGCAGGTCACGGTCAAGGTGGGAATGGTTTCGATGACCTTCCGAAAGAAC 1199
 Db 1121 GAAGGGCATCAGCATTGCCGAGAATGGACAAAACGGGTGATCTTGGAGGGGT 1180
 Qy 1200 TGAGGTGTCCTCCGGTAAT 1220
 Db 1181 TGAGCTCTGACTCATGTATT 1201

RESULT 12
 US-09-070-927A-123
 ; Sequence 123, Application US/09070927A
 ; Patent No. US20020120116A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles A. Kunsch
 ; Patrick J. Dillon
 ; Steven Barash
 ; TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
 ; NUMBER OF SEQUENCES: 982
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/070, 927A
 ; FILING DATE: 04-May-2000
 ; CLASSIFICATION: <Unknown>
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 60/046, 655
 ; FILING DATE: 1997-05-16
 ; APPLICATION NUMBER: 60/044, 031
 ; FILING DATE: 1997-05-06
 ; APPLICATION NUMBER: 60/066, 009
 ; FILING DATE: 1997-11-14
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kenley K. Hoover
 ; REGISTRATION NUMBER: 40, 302
 ; TELECOMMUNICATION INFORMATION:
 ; PHONE: (301) 309-8504
 ; TELEFAX: (301) 309-8512
 ; INFORMATION FOR SEQ ID NO: 123:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 32768 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 123:
 ; US-09-070-927A-123

Qy 172 GACGCACTGGAAACACATCATGGCGACGTGAGGACTGATCCGGAGCTTCCCTGGGATC 231
 Db 9328 GACGAAAGTGGGAATTCTGTTTGTGAGTGGCAGGAACGGAAATTTCCTCAGGAAACGATT 9387
 Qy 232 GGGTGGTTCAACTTCGATTCTGTCCGAAACGGGGATGTTGATGGCACTGCGAGGC 291
 Db 9388 TAACTGGTCACATATCGACACAGTAGTTAATGGTGGACCTTAGATGGTCAAATATGGT 9447
 Qy 292 GTGGTGTGGCCCTTGAGGGTGCCTGGATGCTGGGAATGGCACTGGCATGGCAT 351
 Db 9448 GTCATTGCTTCCATGACAGCTGTCAATATTAGACAATATGGAAACCACGGGT 9507
 Qy 352 CAATTGAGTTCATCGCGATCGTGGAGGCTACGTGAACTGGCATGTTG 411
 Db 9508 TCTTTAGAAGTGCTTCCATGGCTGAAGAAGAAGGGAGTCGTTTCAACGGTTTCTGG 9567
 Qy 412 GCGGGCGGGCCATTGAGGGTTGGTGGAGGAGAAGGGCCCCTGAGCTGGACTCTGGTGTGAG 471
 Db 9568 GCAAGCAAAATGTCGTTGGTAAGCAAAACGGCAAAGTGAAGAAGTAGTAGTCAAC 9627
 Qy 472 GATGGAGTGTCCGTTAGGCAGGGCTACTGCCTTGGCTTAAGGCCGGGAACTGGACTCTTTGGTGTGAG 531
 Db 9628 GAAGGTTAAAAGTTGTTGATGAGATGCACTGCCAAGGGCTTGTATTTTCGAAAGGCAA 9687
 Qy 532 GCTGCAGCCCGCTCCGGGGACCTGCGTGTCTTATCGAACTACACATGGAAAGGA 591
 Db 9688 ACTAGAC-----
 Qy 592 CCGATCCTCGAGCAGGAAATAAGAGATGGAGTTGTAACCTCCATCGTTGGCGTTGCG 651
 Db 9739 ATATTTAGAACCCAAAGCTCCAAGTGGCTTGTGTGTCATAGTATTGTAGGGCAACGT 9798
 Qy 652 GCATTGGGTTGGCTCAAAGGCAGAAGGGGACCAACGCCAACTGGTACCTGACCTG 711
 Db 9799 CGCTACACGGTTAACTTAAAGGCCAAGGCAATTGAGGTACTACGCCAATGAGTTAC 9858
 Qy 712 CGCCAGGATGGCTGACCGCTCTCATGGTACCGGCTAACCGGGTCAACCGGTCTGTCAAC 771
 Db 9859 CGTCATGATGGGGTGTACGGCTTGTCTAAATTTGTGCCAGAACGCCATTGATCGTGCAGAC 9918
 Qy 772 GAGATCGCCGATGGCACAGTGGCTACCGGCCACCTCACAGTGGCCACCTGGGAGGTGG 831
 Db 9919 AACGGTGGTGCCTGGAAAGTGGAAATTACATTGGCAAAAGTCAATTGGAAAGTGG 9978
 Qy 832 AACAGGGTCCGGGGAGGGTGGACTCACACTGGGACTCTGGCTTCTCCAGGGCGGT 951
 Db 10039 TTACGACATTTCACGGTGAATTAGAAGAACGGATGAAGGACAATTGGCTAAATGGCTTT 100398
 Qy 952 GTGGCTGCCGATGTGGATGAATTTCGAATCTCAGGCCGGCTTGTGAGTGGCTCCTACCATG 1011
 Db 10099 TTGAGGATGGCTGCTGGATGGATGAAGGACAATTGGCTCAAGGAATTAGGC 10098
 Qy 10099 TTGAGGATGGCTGCTGGATGGATGAAGGACAATTGGCTCAAGGAATTAGGC 10098
 Db 10119 GTAGAAAGTATTGAAACAAAAGCAGGAAAGGCAAGTGGATGATGGCTATAGT 10218
 Qy 1012 GTGGACGCCACGACTCGATGTTCATGGCCAGGGCTCGGGCTTGTGAGTTCAACACCGGGATATCAGCAGT 1071
 Db 10159 GTAGAAAGTATTGAAACAAAAGCAGGAAAGGCAAGTGGATGATGGCTATAGT 10218
 Qy 1072 GGGGGGCCACGACTCGATGTTCATGGCCAGGGCTCGGGCTTGTGAGTTCAACACCGGGATATCAGCAGT 1131
 Db 10219 GTGGCTGGCATGGCAATTATTGGCCCTCATACGGCTCATACGGCTGCTGCTT 10278
 Qy 1132 CCAAGCGGTGGCTGGGGAGGCCACGCTGGGAATGGACCTTCGATGACCTTCGCG 1191
 Db 10279 CCAAGTATTGGGGATTAGTCATAATCCCTGAGAACGGCATTCAGGATTGGTT 10338
 Qy 1192 AAAGGAACGTGAGGTTGT 1208
 Db 10339 GAAGGTGTCAAGCTTT 10355

RESULT 13
 US-10-329-960-1
 Sequence 1, Application US/10329960
 Publication No. US2003009927A1
 GENERAL INFORMATION:
 APPLICANT: Fleischmann et al.
 TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Frag
 FILE REFERENCE: PB186P1
 CURRENT APPLICATION NUMBER: US/10/329,960
 CURRENT FILING DATE: 2003-01-02
 PRIOR APPLICATION NUMBER: US 09/643,990
 PRIOR FILING DATE: 2000-08-23
 PRIOR APPLICATION NUMBER: US 08/487,429
 PRIOR FILING DATE: 1995-06-07
 PRIOR APPLICATION NUMBER: US 08/426,787
 PRIOR FILING DATE: 1995-04-21
 NUMBER OF SEQ ID NOS: 1
 SOFTWARE: Patentin version 3.1
 SEQ ID NO 1
 LENGTH: 1830121
 TYPE: DNA
 ORGANISM: Haemophilus influenzae
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (4747)..(4747)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (9921)..(9921)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (10150)..(10150)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (29298)..(29298)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (36543)..(36543)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (36551)..(36551)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (36636)..(36636)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (40808)..(40810)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (44416)..(44416)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (44905)..(44905)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (44975)..(44975)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (45593)..(45593)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (45732)..(45732)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (47036)..(47036)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (51334)..(51334)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (51602)..(51602)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (51786)..(51786)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (51805)..(51805)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (55369)..(55369)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (65309)..(65309)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (65313)..(65313)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (65319)..(65319)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (80024)..(80024)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (100091)..(100091)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (102696)..(102696)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (105121)..(105121)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (107248)..(107248)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (117136)..(117136)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (119750)..(119750)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (119924)..(119924)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (120038)..(120038)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature

LOCATION: (121344) .. (121344)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (122167) .. (122167)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (122336) .. (122336)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (131340) .. (131340)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (131360) .. (131360)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (139910) .. (139910)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (140398) .. (140398)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (142750) .. (142750)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (145058) .. (145058)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (145171) .. (145171)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (145942) .. (145942)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (147197) .. (147197)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (150841) .. (150841)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (152500) .. (152500)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (152530) .. (152530)

Qy 208 ACTGATCCGGAGCTTCTGCGATGCCGTGGTTCAACACTTCGATTCGATGCCAAACGGC 267
 Db 608840 AAAGAAGATTTTAACCTGCGATTCGATGCATTGGATCAATATTGATACTGTTGTGAATGCT 608899
 Qy 268 CGGATGTTGATGGCACTGCGGTGGCTGCCCTTGAGGCTGGCTGGATGCTG 327
 Db 608900 GTTAAATTGATGGTCTTGGATCTGGTGTGGTTGGAATTCTTTGCAATTGTTGTGT 608959
 Qy 328 GAGAACGGCTACGTTGAGTTCATCGCGATCGTGGAGGGAGGG 387
 Db 608960 GAACAGAAATTAGCTCGTTAGAATTGATAATTTTACTTGTGAAGAATATCG 609019
 Qy 388 GCCCGCTTACGGAGTGGCATGTTGGGGCCATTGCGAGGTTGGCTGGCGACAGG 447
 Db 609020 AGTAGATTAAATTGGTACATTTGCTAAGTTAGTGTGGCATAGTAATCAAGAA 609079
 Qy 448 GAACTGGACTCTTGTGATGAGGATGGAGTGGCTACTGGCTTC 507
 Db 609080 AAATTAAGTTCAATTACGTGATAAACAGGAAGGGTTATCAGAAAGCTATGGCTGAAGTA 609139
 Qy 508 GGCTTGAAAGCCGGGCAACTGCAGGCTGAGCCCGCTCGCGGGAACCTGGCTTT 567
 Db 609140 GGAATGAATTAAATTGGTTAAATCAAGCAAACGTGATGCAAAGGAATTAAATGTTT 609199
 Qy 568 ATCGAACTACACATGGAAACAAGGACCGATCCTCGAGCAGGAAATAGAGATCGGAGTT 627
 Db 609200 TTGAACTCATATAGAACAGGCCCTCGTTAGAAATGAAAGAAAACAATTAGGTGTT 609259
 Qy 628 GAAACCTCATCGTGGCTTCGCGATTGGGGTTCGCGTCAAAGGCAGAAGCGACCAC 687
 Db 609260 GTGACAGGGATTGCTGCCCGATCTGGCAATTGTTAAAGGACAAGGGATCAT 609319
 Qy 688 GCCGGCACAAACCCCCATGCAACCTTGCCAGGATGCGCTGGTACCCGCCGCTCTCATGGTG 747
 Db 609320 TCAGGAGAACAGCAATGCAATGCAATGCTCATGATGCTATGAGGATCTGAGTTATCA 609379
 Qy 748 AGGGAGGTCAACGGTTCGTCAACGAGATGCCGATGCCGACAGTGGCTACCGTTGGCCAC 807
 Db 609380 CTTGGCTATGAGGAGCTGCTATCAAGCTGGACATTCAAGCTGCTAGTAGTTAAT 609439
 Qy 808 CTACAGTGGCCCGTGGAGGCAACCGAGTCCGGGAGGTGGACTTCACACTGGAC 867
 Db 609440 ATTACAGCTAACCGAGGAACTGAAATGAAATGTTGCAATTATTAGTAGAT 609499
 Qy 868 CTGGAGGTCAACGGTTCTCGTCAACGAGATGCCGATGCCGACAGTGGCTACCGTTGGCCAC 927
 Db 609500 ATTCAAGGTACATGTACAAGCTGCTATGAGGATTTGAAATTATTGAAATTACAGCTGCTAGTAGTTAAT 609559
 Qy 928 GGGGAGGTGGCTCCGAGGGTGGATGTTGGATGAATTTTCAATCTCAGC 987
 Db 609560 AGTAAAGTTCAAAAAAGAGGATTATTAAATTGAGATTCTGTTGAAATTAGCTGAAACAGGAAATT 609619
 Qy 988 CCAATAATATTACCTGAAAAAATGGTAAACCCAATAGCTGAAACAGCACATTCCCTGGT 609679
 Db 609620 CTCACACACCGGGATATCAGCAGTGGGGGGCCACGACTCGATGTTCATCGCCCCGGTC 1107
 Query Match 9.8%; Score 121.2; DB 15; Length 1830121;
 Best Local Similarity 43.8%; Pred. No. 2.7e-25;
 Matches 525; Conservative 0; Mismatches 673; Indels 0; Gaps 0;

Qy 28 CGCATTTGAGAAAGAGATCCGGAGCTCCGGTCTCGGCAAGAGGCCGGTGTAC 87
 Db 608660 CGTGTTCAAAATTAAATTGAAAAAAACTGGTTTTATTTCATCAGTACCAAAACGGCTGACT 608719
 Qy 88 CGGCTGACCTACACTCCAGAGCATGCCCGGGAAACGCTCATGGGGCTATGAAA 147
 Db 608720 CGCTTAGCTTTACAGGGAGGATGAAAAGCCCATAATAATTGAAATTATGTAAA 608779
 Qy 148 CGGGCCGCCTTGAAGACGCACTCGGAAACATCATCGGGCGACGTGAAGGC 207
 Db 608780 GAATATGATTGGAAAGATAAGGGCAGGTTTACAAAAAGTTATATTGG 609857

RESULT 14
 US-10-329-670-1
 ; Sequence 1, Application US/10329670

Publication No. US20040018503A1
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd Genome, Fragments Thereof, and Uses Thereof
FILE REFERENCE: PB186P1
CURRENT APPLICATION NUMBER: US/10/329,670
CURRENT FILING DATE: 2002-12-24
PRIOR APPLICATION NUMBER: US 09/643,990
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: US 08/487,429
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
PRIOR FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn version 3.1
SEQ ID NO: 1
LENGTH: 1830121
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (12593)..(12593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc feature
LOCATION: (45593)..(45593)

LOCATION: (122167)..(122167)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (122336)..(122336)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (131340)..(131340)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (131360)..(131360)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (139910)..(139910)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (140398)..(140398)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (142750)..(142750)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (145058)..(145058)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (145171)..(145171)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (145942)..(145942)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (147197)..(147197)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (150841)..(150841)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (152500)..(152500)
 OTHER INFORMATION: n equals a, t, g or c
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (152530)..(152530)

Query Match 9.8%; Score 121.2; DB 16; Length 1830121;
 Best Local Similarity 43.8%; Pred. No. 2.7e-25;
 Matches 525; Conservative 0; Mismatches 673; Indels 0; Gaps 0;

Qy 268 CGGATGGTGCCTTGAGCGTTCTGGCCGACGTGGCTGCCCTTGAGGCTGCCCTTGAGGCTGATGCTG 327
 Db 608900 GGTAATTGATGGTCCTTGGGATCTCTGCTGGTTGGAAATTCTTCAATTGGAAATTGGTGTGTTG 608959

Qy 328 GAGAACGGCTACGTGAATCGGCATCCATTGAGTTCATCGCGATCGTGGAGGGAGGAAGGG 387
 Db 608960 GACAGAAATTTCAGACTCGTTATCCTTAGAATTGATAATTGTTACTTGTGAAGAATTCG 609019

Qy 388 GCCCGCTTCAGCAGTGGCATGTTGGGCCCCCATGGAGGGTGGCTGCCGACAGG 447
 Db 609020 AGTAGATTAAATTGGTACATGGTAGTAAAGTTATGTGGCATAGTAATTCAAGAA 609079

Qy 448 GAACTGGACTCTTGGTTGATGAGGATGGAGTGGCTACTGGCTTC 507
 Db 609080 AAATAAGGTCAATTACGTGATAAACAGGAAGGGTTATCAGAAGCTATGGCTGAAGTA 609139

Qy 508 GCCTTGAGCCGGCAACTGCAGGCTGAGCCCGCTCAGGGGAGCTGGCTGCTTT 567
 Db 609140 GGAATGAAATTAAATTGGTTAACTCAAGGAAAACGTGATGCAAAAGGAATTAAATGTTT 609199

Qy 568 ATCGAACTACACATTGAAACAAGGACCGATCCTCGAGCAGGAAATTAGAGATCGGAGTT 627
 Db 609200 TTGAACTTCATATGAAACAAGGCCCTCGTTAGAGAAAGAAAACAATTAGGTGTT 609259

Qy 628 GAAACCTCCATCGTTGGCGCATGGTGGCGTCAAAGGCAGAAGGCACAC 687
 Db 609260 GTGACAGGTATTGCTGCCCGATCTGTGCAATTGGTAAATTAAAGGACAAGGGATCAT 609319

Qy 688 GCGGGCACACCCCCATGCACCTGGCCAGGATGGCTTACCCGGCTCATGGTG 747
 Db 609320 TCAGGAGAACAGCAAATGCATTATCGTCAATGATGCTTACAGTTAGAGGATCTGAGTTATCA 609379

Qy 748 AGGGAGGTCAACCGGTTCGTCAAAGGAGATGGCACAGTGGCTACCCGGTGGACTTCAACACTGGAC 807
 Db 609380 CTTGCTTATGAGGAGCTGGTGTGCTATTCAGTTAGCTACAGTAGTTAT 609439

Qy 808 CTACACAGTGGCCCCGGGAGCAACCAGGTCGGTGGTATTCAGTTAGCTACAGTAGTTAT 609440
 Db 609440 ATTACAGCTAAACCGGAGTAATGAAATGTTGTGCCAGGATATTAGTAGAT 609449

Qy 868 CGGGAGGTGGCTTCCCGCATGGAGGTGCTGATGACCCGGTTGGCTGATGGCTCATGGTC 927
 Db 609500 AGTAAAGTTTCAGAAAAAGGGATTATAATTGAGTTACAACATGTACAAGCTAGAT 609559

Qy 928 CGGGAGGTGGCTTCCCGCATGGAGGTGCTGATGACCCGGTTGGCTGATGGCTCATGGTC 987
 Db 609620 CAAATAATTACCTGAAATAATGGTAAACCAAAATAGCTGAAACAGCACATTCCCTGGT 609619

Qy 988 CGGGAGGTGGCTTCCCGCATGGAGGTGCTGATGACCCGGTTGGCTGATGGCTCATGGTC 1047
 Db 609680 TATTCTTACGAAATAATGCCAAGTGGCAGGACATGTCATGGCAATGCAATTGGCAACACTT 609679

Qy 1048 TTCACACACCCGGATTCAGCAGTGGCTGGCCACGACTCGATGTTCATGCCAGGTC 1107
 Db 609740 TGCCCAACAGGTATGATTTCATCTCATTGGAAATTAGTCATAACCCTCTTGAA 609799

Qy 1168 TGGACCGATTTCGATGACCTTCGAAAGGAACGTGAGGTGTCCCTCGGGTAATGAAGG 1225
 Db 609800 TTACTGATTGAAAGATATAGGGCAGGAAATTAAAGTTTACAAAAGTTTACATATTGTTG 609857

RESULT 15
 US-10-156-761-3482 ; Sequence 3482, Application US/10156761
 ; Publication No. US20030119018A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OMURA, SATOSHI
 ; APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN
 APPLICANT: HORIKAWA, HIROSHI
 APPLICANT: SHIBA, TADAYOSHI
 APPLICANT: SAKAKI, YOSHIIYUKI
 APPLICANT: HATTORI, MASAHIRO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-262
 CURRENT APPLICATION NUMBER: US/10/156,761
 CURRENT FILING DATE: 2002-05-29
 PRIOR APPLICATION NUMBER: JP 2001-204089
 PRIOR FILING DATE: 2001-05-30
 PRIOR APPLICATION NUMBER: JP 2001-272697
 PRIOR FILING DATE: 2001-08-02
 NUMBER OF SEQ ID NOS: 15109
 SEQ ID NO: 3482
 LENGTH: 1239
 TYPE: DNA
 ORGANISM: Streptomyces avermitilis
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)..(1239)
 US-10-156-761-3482

Query Match 9.1%; Score 112.8; DB 15; Length 1239;
 Best Local Similarity 48.2%; Pred. No. 1.8e-23;
 Matches 352; Conservative 0; Mismatches 372; Indels 6; Gaps 1;

Qy 238 GGTTCACTTGATTCTGTCGGAAACCGGGGATGTTGATGGACTGGAAGGGTGGGTG 297
 Db 217 GGCTCTCACCTGGACTCCGTCCCCGACGGGGGGCCTTCGACGCCCTGGGTGGT 276

Qy 298 TGCGCCCTTGAGGGCTGCCGGGTGATGGCTGGAGAACGGCTACGGTAATCCATT 357
 Db 277 TCCTCCTCGCCGCCCTCGACGAACCTCCGGCGCCAGGTACGGCTGCCAAGGCCCTC 336

Qy 358 GAGTTCATCGCATCGTGGAGGAAGGGGGCCCTTCAGCAGTGGCATGGGGGC 417
 Db 337 GCCTCTGTGAACCTCGGGACAGCTGGGGACAGGGGGCCCTCGGCTCGGTGGTCC 396

Qy 418 CGGCCATTGCGAGGGTTGGCAGCGAACGGAAACTGGACTCTTGGTTGAGGGATGGA 477
 Db 397 CGCTCGGCCGGACAGCTACCGTCAAGCAGGGCCACCGGGTACCGAACGGCC 456

Qy 478 GTGTCGTTAGGAGGGCTACTGCCTTCGGCTTGAAGCGGGGAACACTGAGGCTGCA 537
 Db 457 GTCACCCCTCCGCAAGGCCATGGAGGGCCATGGGGCTGGCTTATCGAACTACACATTGAAACAGGAC 516

Qy 538 GCGCGCTCCGGGGGACCTGGCTGGCTTGTGGAGGGCTGGCTGGCTGGCTGG 597
 Db 517 CCCGAGGGCTGCCGTATGGCGCTTGTCGAAGCTGGCATGGCAACTGGCTGGCC 576

Qy 598 CTCGAGCAGGCAAATAGAGATGGAGTGGAGGTGTAACCTCCATGGCCATGGCACTGGCTGG 657
 Db 577 CTGGAACCTCAGGGTGAACGGATCGGCATGGCCATGGCCAGGGCCATGGCC 636

Qy 658 CGGGTGTGGTCAAGGGCAGAAGGGGACCAACGGGGCAACACCCCATGGCACTGGCC 717
 Db 637 CGCTTCGACTTCGGGGCGAGGGCAACACGGGGACGGCCCTGGGGACGCC 696

Qy 718 GATGGCGCTGGTACCGGGCTCATGGTGAGGGGGCACCTCACAGTGGCCGGGGTCAACCGGTTCTGGTCAACGGAGTC 777
 Db 697 GACCCGATGCTGTGTACGGGAGACGGTCCCTGGGGACGGCTGGCCGG-----AGAGGCCGA 750

Qy 778 GCGGATGGCACAGGGGACTTACACTGGACCTCACAGTGGCCGGGGTCAACCGGTTCTGGTCAACGGAGTC 837
 Db 751 CTCGCCGGTGGCCACCTCGGCAAGATGCGTCAAGATGCGTCAACGGGGTCAAGGCC 810

Qy 838 GTCCGGGGAGGGGACTTACACTGGACCTCACAGTGGCCGGGGTCAACCGGTTCTGGTCAACGGAGTC 897
 Db 811 ATTCCCTCCCTCGTCCGGGGTGGCTGACTCCGGCCGACCGGGGACCCCTCGAC 870

Qy 898 GTGCTGATGCGACGCCATCTGGTCATGGTGGGGTCAACGGGGTCAACGGGGTCAAGGCC 957

This Page Blank (uspto)